THE INVERSION OF THE BOLTZMANN TRANSPORT EQUATION WITH MEDICAL OPTICAL TOMOGRAPHY APPLICATIONS

by

Victoria Barnard, MPHYS

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy and the Diploma of Imperial College

Imperial College
London

Computational Physics and Geophysics Group
Department of Earth Sciences and Engineering
Imperial College of Science, Technology and Medicine
University of London

November, 2006

Publications as a Result of this Work

V. L. Barnard, C. C. Pain, A. G. Buchan, M. D. Eaton, and A. J. H. Goddard, *A Finite Element Optical Imaging Method for Optically Thick and Optically Thin Media*, Submitted to Inverse Problems, (2006)

V. L. Barnard, C. C. Pain, A. G. Buchan, M. D. Eaton, and A. J. H. Goddard, *A Finite Element Optical Imaging Method Using Time Dependent Information*, Submitted to Journal of Qunatitative Spectroscopy and Radiative Transfer, (2006)

V. L. Barnard, C. C. Pain, A. G. Buchan, M. D. Eaton, and A. J. H. Goddard, *Optical Imaging with Voids using Transport Theory*, Submitted to Physics in Medicine and Biology, (2006)

Abstract

A new multi-dimensional inversion technique has been developed here to describe the transport of neutral particles inside a host medium. This thesis focuses on applications relating to the transport of near infrared photons during medical optical tomography. The nature of the inversion problem is to reconstruct cross-sectional images of the optical properties of highly scattering biological tissue from source measurements and detector readings performed on the surface of the domain. The transport of optical radiation is described according to the one-speed time-dependent Boltzmann transport equation. A Streamline Upwinded Petrov-Galerkin (SUPG) finite element method is used to spatially discretise the Boltzmann transport equation, a Discontinuous Galerkin (DG) method is used to discretise the time domain, and spherical harmonic basis functions are used to represent the angle of photon travel. SUPG methods form a robust coupling between time and space dimensions. A stabilisation term, designed specifically for inverse problems, is incorporated into the discretisation to ensure an optimal modelling method for both optically thick and optically thin media. We demonstrate the capabilities of inversion using both steady state and time dependent information and draw comparisons between full radiation transport theory and diffusion theory.

A least squares functional optimisation method is used to solve the inverse problem. An error functional representing the forward model misfit measured at the detectors is minimised using a gradient based, least squares optimisation scheme and is obtained by differentiating the discrete forward model. This results in the reconstruction of the spatial distribution of scattering and absorption coefficients inside the domain. Photon propagation in layered tissue environments can produce non-unique solutions to the inverse problem when insufficient data is supplied to the model. Special attention is therefore given to the use of model covariance matrices and data weighting to assist the inversion process to arrive at a physically plausible result.

Acknowledgements

I would like to thank my supervisors, Professor Christopher Pain and Professor Anthony Goddard, for sharing their knowledge and providing support and encouragement throughout my PhD. I am also grateful to Professor Mike Williams for his helpful advice on radiation transport physics and Dr Fangxin Fang for patiently sharing her inversion expertise during early stages of the research. I owe an enormous thank you to Mr Andrew Buchan who painstakingly developed RADIANT's fast solvers. Without his help, my PhD would have surely stagnated.

I would like to thank Mr Andrew Buchan and each of my family members for their personal support. I am particularly grateful to my fiancè, Blair Hooley and my parents, Roy and Carol Barnard, for both their support and financial help.

This research would not have been possible without the funding supplied by EPSRC. I would like to thank Dr Cassiano de Oliveira for his help in obtaining this funding and Professor Anthony Goddard for providing a small amount of additional funding after the EPSRC studentship had run out.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms and Abbreviations

CerebroSpinal Fluid	CSF
(X-ray) Computed Tomography	CT
Control Volume	CV
Discontinuous Galerkin	DG
Discrete Ordinates (method)	S_N
Finite Difference Method	FDM
Finite Element Method	FEM
IntraVeneous Pyelography	IVP
Kidney Urethra and Bladder (scans)	KUB
Magnetic Resonance Imaging	MRI
Optical Tomography	OT
Positron Emission Tomography	PET
Preconditioned Conjugate Gradient (solver)	PCG
RADIAtion Non-oscillatory Transport	RADIANT
Self Adjoint Angular Flux	SAAF
Single Photon Emission Tomography	SPECT
Spherical Harmonics (method)	P_N
Streamline Upwind Petrov-Galerkin	SUPG

Greek Characters

Step Length	α
Cosine spherical harmonic (scattering) expansion coefficient	$\alpha^c_{l,m}$
Sine spherical harmonic (scattering) expansion coefficient	$\alpha^s_{l,m}$

Dirac delta	δ
Denotes a change	Δ
Time domain of time level τ	ΔT^{τ}
Residual vector	ϵ
Lagrangian multiplier associated with absorption	γ_a
Lagrangian multiplier associated with scattering	γ_{s0}
Boundary of solution domain	Γ
Infinity	∞
Interpolation length	κ
Conditioning factor	λ
Eigenvalues of scattering/removal operator	$\mathbf{\Lambda}_H$
Eigenvalues associated with Riemann boundary conditions	$\mathbf{\Lambda}_s$
Co-latitude from the z axis	μ
Energy dependent absorption coefficient	$\mu_a(\mathbf{r}, E)$
Energy independent absorption coefficient	$\mu_a(\mathbf{r})$
Energy dependent scattering coefficient	$\mu_s(\mathbf{r}, E)$
Energy independent scattering coefficient	$\mu_s(\mathbf{r})$
Energy dependent reduced scattering coefficient	$\mu_s^{'}(\mathbf{r}, E)$
Energy independent reduced scattering coefficient	$\mu_s^{'}(\mathbf{r})$
Energy dependent differential scattering cross-section	$\mu_{\mathrm{s}}(\mathbf{r}, \mathrm{E}' \rightarrow \mathrm{E}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega})$
Energy independent differential scattering cross-section	$\mu_{\mathrm{s}}(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega})$
Angularly discretised scattering moment	$\mu_{\mathrm{sl}}(\mathbf{r})$
Energy dependent total attenuation coefficient	$\mu_t(\mathbf{r}, E)$
Energy independent total attenuation coefficient	$\mu_t(\mathbf{r})$
Annealing parameter associated with absorption regularisation	ν_a
Annealing parameter associated with scattering regularisation	ν_{s0}
Azimuthal angle	ω
Angular direction	$\mathbf{\Omega}$
Component of angular direction vector in the $k = x, y, z$ direction	Ω_k

Component of angular direction vector in the x direction	Ω_x
Component of angular direction vector in the y direction	Ω_y
Component of angular direction vector in the z direction	Ω_z
Scalar constant	ϕ
Pi	π
Angular flux	$\psi(\mathbf{r}, \mathbf{\Omega}, \mathrm{E}, t)$
Energy independent angular flux	$\psi(\mathbf{r}, \mathbf{\Omega}, t)$
Energy and time independent angular flux	$\psi(\mathbf{r}, \mathbf{\Omega})$
Angularly discretised flux	$\mathbf{\Psi}(\mathbf{r}, t)$
Angularly discretised flux	$\boldsymbol{\psi}(\mathbf{r}, t)$
Angularly discretised time independent flux	$\mathbf{\Psi}(\mathbf{r})$
Angularly discretised time independent flux	$\boldsymbol{\psi}(\mathbf{r})$
Moments for the angularly discretised scalar flux	$\mathbf{\Psi}_k$
Moments for the angularly discretised scalar flux at time level τ	$\mathbf{\Psi}_j^{\tau}$
Angularly discretised flux as an infinite set of spherical harmonics	$\Psi_{l,m}^{\mathcal{R},\mathcal{C}}(\mathbf{r}, t)$
Angularly discretised flux as an infinite set of spherical harmonics	$\Psi_{l,m}(\mathbf{r}, t)$
Cosine spherical harmonic (flux) expansion coefficient	$\Psi_{l,m}^{c}(\mathbf{r}, t)$
Sine spherical harmonic (flux) expansion coefficient	$\Psi_{l,m}^{s}(\mathbf{r}, t)$
Microscopic absorption cross-section	$\sigma_{\mathrm{a}}(E)$
Microscopic scattering cross-section	$\sigma_{\mathrm{s}}(E)$
Reduced microscopic scattering cross-section	$\sigma_{\mathrm{s}}^{'}(E)$
Energy Dependent scattering phase function	$\sigma_{\mathrm{s}}(\mathrm{E}' \rightarrow \mathrm{E}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega})$
Time level	τ
Polar angle	θ

Latin Characters

Angular jacobian matrix	$\mathbf{A}$

Angular jacobian matrix in the $k = x, y, z$ direction	$\mathbf{A}_k$
Matrix equal to $\mathbf{A} \cdot \mathbf{n}$	$\mathbf{A}_\mathrm{s}$
Angular mass matrix	$\mathbf{A}_t$
Angular jacobian matrix in the x direction	$\mathbf{A}_x$
Vector of space-time angular jacobian matrices	$\mathbf{A}_{xt}$
Angular jacobian matrix in the y direction	$\mathbf{A}_y$
Angular jacobian matrix in the z direction	$\mathbf{A}_z$
Gradient vector	$\mathbf{a}$
Boundary (source) vector	$\mathbf{b}$
Hessian matrix	$\mathbf{B}$
Preconditioning matrix	$\hat{\mathbf{B}}$
Magnitude of source strength	C
Diffusion coefficient	$D(\mathbf{r})$
Energy variable	E
Matrix used to represent the linearised Boltzmann transport equation	$\mathbf{E}$
Error functional	F
Data misfit	F_d
Spatial regularisation functional	F_r
Deviation from the starting model regularisation functional	F_v
Differential phase function	$f(E' \rightarrow E, \mathbf{\Omega}' \rightarrow \mathbf{\Omega})$
Anisotropy factor	g
Matrix of modified eigenvalues	$\mathbf{G}(\mathbf{\Lambda}_H)$
Matrix used to simplify the description of $\mathbf{P}$	$\mathbf{g}(h\mathbf{\Lambda}_H)$
Chraracteristic element size	h
Angularly and spatially discretised scattering/removal operator	$\boldsymbol{\mathcal{H}}(\mathbf{r})$
Angularly discretised scattering/removal operator	$\mathbf{H}(\mathbf{r})$
Expansion coefficients for the spatially discretised material properties	$\mathbf{H}_j$
Identity Matrix	$\mathbf{I}$
Jacobian Matrix	$\mathbf{J}$

Approximate Hessian matrix	$\mathbf{J^T W J \Delta q}$
Regularisation constraint	k
Matrix of regularisation constraints	$\mathbf{K}$
Covariance matrix	$\mathbf{K}^{-1}$
Truncation length	L
Detector length over a which an interpolation is performed	l_r
Source length over a which an interpolation is performed	l_s
Control variable vector (material properties)	$\mathbf{m}$
Control variable vector (absorption material properties)	$\mathbf{m}_a$
Absorption material properties associated with previous update	$\mathbf{m}_{a_{obs}}$
Control variable vector (scattering material properties)	$\mathbf{m}_{s0}$
Scattering material properties associated with previous update	$\mathbf{m}_{s0_{obs}}$
Material properties associated with the previous update	$\mathbf{m}_{old}$
Updated material properties	$\mathbf{m}_{new}$
Lumped mass matrix	$\mathbf{M}_{H_L}$
Finite element basis associated with node ν	$M_\nu(\mathbf{\Omega})$
Finite element basis associated with node ρ	$M_\rho(\mathbf{\Omega})$
Normal boundary vector	$\mathbf{n}$
Normalisation coefficient	n_s
Normal component in the direction of time	n_t
Normal component in the x direction	n_x
Normal component in the y direction	n_y
Normal component in the z direction	n_z
Matrix of space-time finite element basis functions	$\mathbf{N}(\mathbf{r}, t)$
Matrix of spatial finite element basis functions associated with materials	$\mathbf{N}_H(\mathbf{r})$
Matrix of spatial finite element basis functions	$\mathbf{N}(\mathbf{r})$
Matrix of spatial finite element basis functions associated with sources	$\mathbf{N}_s(\mathbf{r})$
Finite element variation in time	$\mathbf{N}^\tau(\mathbf{r}, t)$
Finite element variation in time associated with source terms	$\mathbf{N}_s^\tau(\mathbf{r}, t)$

SUPG Stabilisation matrix	$\mathbf{P}$
Legendre Polynomial	$P_l(\mu)$
Associated Legendre Polynomial	$P_{l,m}(\mu)$
Arbitrary perturbation vector	$\mathbf{q}$
Sturctural vector contained inside global matrix $\mathbf{E}$	$\mathbf{Q}$
Energy independent source contributions due to scattered particles	$q_s(\mathbf{r}, \mathbf{\Omega}, t)$
Matrix of right eigenvectors	$\mathbf{R}_H$
Matrix of right eigenvectors associated with Riemann boundary conditions	$\mathbf{R}_s$
Matrix of left eigenvectors	$\mathbf{R}_H^{-1}$
Matrix of left eigenvectors associated with Riemann boundary conditions	$\mathbf{R}_s^{-1}$
Position vector	$\mathbf{r}$
Surface position vector	$\mathbf{r}_s$
Source	$s(\mathbf{r}, \mathbf{\Omega}, \mathrm{E}, t)$
Energy independent source	$s(\mathbf{r}, \mathbf{\Omega}, t)$
Energy and time independent source	$s(\mathbf{r}, \mathbf{\Omega})$
Angularly discretised source	$\mathbf{S}(\mathbf{r}, t)$
Angularly discretised source	$\mathbf{s}(\mathbf{r}, t)$
Angularly discretised time independent source	$\mathbf{S}(\mathbf{r})$
Angularly discretised time independent source	$\mathbf{s}(\mathbf{r})$
Moments for the angularly discretised source	$\mathbf{S}_k$
Moments for the angularly discretised source at time level τ	$\mathbf{S}_j^\tau$
Time	T
Time variable	t
Sturctural vector contained inside global matrix $\mathbf{E}$	$\mathbf{U}$
Volume	V
Diagonal conditioning matrix	$\mathbf{V}$
Particle velocity	v
Arbitrary perturbation vector	$\mathbf{v}$
Matrix containing Gaussian weights	$\mathbf{W}$

Individual Gaussian weight	w
Quadrature weight function	$w_\nu(\mathbf{\Omega})$
Position vector	$\mathbf{x}$
Position vector	$\mathbf{X}$
x direction	*x*
y direction	*y*
Spherical harmonic containing real and complex terms	$\mathrm{Y}_{l,m}^{\mathcal{R},\mathcal{C}}(\mathbf{\Omega})$
Cosine real orthonormal spherical harmonic	$\mathrm{Y}_{l,m}^{\mathrm{c}}(\mathbf{\Omega})$
Sine real orthonormal spherical harmonic	$\mathrm{Y}_{l,m}^{\mathrm{s}}(\mathbf{\Omega})$
Compex spherical harmonic	$\mathrm{Y}_{l,m}^{\mathcal{C}}(\mathbf{\Omega})$
Real orthornomal spherical harmonic	$\mathrm{Y}_{l,m}^{\mathcal{R}}(\mathbf{\Omega})$
Cosine real orthonormal spherical harmonic	$\mathrm{Y}_{l,m}^{\mathrm{c}}(\mathbf{\Omega})$
Sine real orthonormal spherical harmonic	$\mathrm{Y}_{l,m}^{\mathrm{s}}(\mathbf{\Omega})$
Compex spherical harmonic	$\mathrm{Y}_{l,m}^{\mathcal{C}}(\mathbf{\Omega})$
z direction	*z*

Calligraphic Characters

Atomic number/ Particle Density	$\mathcal{A}(\mathbf{r})$
Scattering/removal operator	$\mathcal{H}$
Photon current density	$\mathscr{J}(\mathbf{r},t)$
Jacobian matrix at a given node	$\mathcal{J}$
Number of finite element basis functions associated with angular moments	$\mathcal{M}$
Number of space-time finite element basis functions associated with flux and sources	$\mathcal{N}$
Number of space-time finite element basis functions associated with sources	$\mathcal{N}_s$
Number of space-time finite element basis functions associated with materials	$\mathcal{N}_H$
Number of space-time finite element basis functions associated with time	$\mathcal{N}_t$
SUPG residual vector	$\mathscr{R}$
Total number of detectors	$\mathcal{R}$
Total number of sources	$\mathcal{S}$

Total number of time levels	$\mathcal{T}$

Mathematical Operators and Symbols

Gradient operator	$\boldsymbol{\nabla}$
Gradient operator associated with space-time	$\boldsymbol{\nabla}_{xt}$

Superscipts and Subscripts

Subscript to denote an initial component	0
Superscript to denote complex term	$\mathcal{C}$
General subscript to denote a component or an imaginary number	i
Subscript denotes incoming	in
General subscript to denote a component	j
Subscript to denote x-, y-, and z-direction or a general component	k
General subscript to denote a component or an imaginary number	μ
General superscript to denote a nodal position	ν
Subscript denotes outgoing	out
General superscript to denote a nodal position	ρ
Supercript to denote a change in particle direction	$\prime$
Superscript to denote real orthonormal term	$\mathcal{R}$
Detector component	r
Source component and boundary component	s
Time component	t
Superscript to denote a matrix transpose	T
Subscript to denote x-direction	x
Subscript to denote y-direction	y
Subscript to denote z-direction	z
Superscript to denote a complex conjugate or an adjoint variable	*

Chapter 1

Introduction

Optical tomography is a non-invasive and relatively inexpensive modality capable of functional imaging [2, 3, 4, 5]. During optical tomography near infrared light, of wavelength 650nm-900nm, is guided, on to the surface of a subject using a series of fibre optic sources. While inside the medium, the near infrared radiation undergoes a series of absorption and scattering events before re-emerging through the surface where it is received by detection apparatus [6, 7]. A reconstruction of the spatial distribution of optical properties (in this work - scattering and absorption coefficients) contained inside the medium is attempted from the transmitted and/or reflected intensities received by the detectors. Subject areas that benefit from such technology include medical physics, oceanography, atmospheric science, astrophysics, and neutron physics [8, 9, 10, 11].

The field of biomedical optics [12, 13] has greatly benefited from tomographic technology in diagnosing and monitoring the treatment of disease. Pathological conditions often provide a contrast in tissue optical properties when compared to healthy tissue [14]. These contrasts can be exploited to provide (three dimensional) qualitative and quantitative optical images. Several fields of medicine have particularly benefited from optical techniques. Optical imaging has been used to functionally image brain activity [12, 15, 16, 17], monitor blood oxygenation [18, 19, 20, 21], and to detect cerebral haemorrhaging [6, 22], osteoarthritis, rheumatoid arthritis [23, 24], and Alzheimer's disease [25]. In cancer and tumour research, optical tomography could be used in the diagno-

sis and monitoring of breast [26], prostate and brain tumours along with skin and breast cancers [27, 28, 29, 30, 31]. The non-invasive properties of optical radiation together with optical tomography's relative affordability compared to most conventional imaging technology make it a desirable method to develop and utilise as a mainstream imaging procedure.

The quality of the optical property reconstruction is strongly dependent on the accuracy of the forward model. Since the transport of near infrared photons in biological media is generally regarded as diffuse, many researchers use the diffusion approximation [32, 33, 34] to approximate the more generally applicable Boltzmann transport equation [3, 4, 26, 35, 36]. However, the diffusion approximation fails in regions where the absorption coefficient is comparable to the scattering coefficient, and also in regions where both scattering and absorption are either very low or negligible (void-like regions) [36, 37]. The clear, virtually transparent, layer of cerebrospinal fluid (CSF) that surrounds the human brain makes brain imaging particularly problematic using diffusion theory. Regions that do not comply with the constraints of the diffusion approximation must be accounted for if optical tomography should develop into a conventional imaging tool. Numerical methods that can tolerate tissues that do not comply with the constraints of the diffusion approximation include Monte Carlo [38, 39] and full radiation transport models that rely on the Boltzmann transport equation [40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51]. Although Monte Carlo methods provide correct solutions the simulations are often too slow to be used for realistic geometries such as those required in brain and breast imaging. Radiation transport models that solve the Boltzmann transport equation provide a suitable alternative to diffusion theory but are often difficult to solve and require much larger computational time scales to perform an image reconstruction. As such, many research groups still rely on either diffusion theory or radiation-diffusion hybrid models to perform tomographic image reconstructions [52, 53, 54].

Incorporating time dependent information into the imaging domain can also be used to improve the accuracy of the forward model [50, 51]. However, in many areas of optical imaging, scattering dominant light transport often makes it difficult or even impossible to observe an accurate temporal signal [55, 56, 57]. Factors that reduce signal quality include, noise, the small size of the time window in which photons are observed, and

the relatively small number of photons observed at the detectors (e.g. when imaging the neonatal brain). A finite element, Streamline Upwinded Petrov-Galerkin (SUPG) technique [58, 59, 60, 61, 62] has been used here to discretise the Boltzmann transport equation in space and time. Angular discretisation is achieved using spherical harmonics. SUPG methods are accurate and robust across all radiation transport regimes from optically thick to thin media and provide an accurate coupling between the time and space dimensions. For inversion regimes, the SUPG method is designed to yield a set of discrete equations that can be differentiated with respect to an object's material properties. This allows gradients to be formed as part of the inversion procedure. The SUPG time stepping method uses a Discontinuous Galerkin (DG) discretisation of the time domain that has the ability to use large time steps while still capturing the essential features of the time dependent signal.

Scattering dominant light transport [63, 64, 65] is a limiting factor in the development of optical tomography. During most conventional medical imaging techniques (such as CT and SPECT) scattering is neglected and the path of the detected photon is assumed to be a straight line between a chosen source-detector pair. For low scattering environments the reconstruction problem can therefore be formulated, as a system of linear equations, directly derived from the Boltzmann transport equation. For strongly scattering environments, such as those posed in near infrared optical tomography, the progressive influence of scattering with distance makes the scattering problem complex and non-linear [66, 67]. The solution therefore does not only involve the calculation of optical properties but must also take in to account the intensity and direction of the scattered photons while inside the host medium. Consequently there is not a generally applicable direct method for imaging highly scattering phenomena, and instead, non-linear, iterative inversion methods are applied to evaluate optical properties [3, 4, 26, 35, 40, 41, 42, 68, 69, 70]. Such techniques include the first order methods such as the non-linear conjugate gradient method and techniques that use second order methods such as the Levenberg-Marquardt method [71, 72, 73].

A multi-dimensional inversion scheme is developed in this thesis to describe the transport of photons during near infrared optical tomography. The technique enables the reconstruction of optical property distribution in two and three dimensions for arbitrary model

geometries and for arbitrary source and detector positions along the boundary of the domain. The inverse problem is posed as a functional optimisation problem [41, 69, 74, 75, 76]. Functional optimisation techniques employ a forward model that provides detector readings based on estimates of the distribution of optical properties inside the medium. The predicted detector readings are compared alongside observed data and are quantified using an error functional. The functional is minimised by iteratively updating the trial estimates and performing new forward calculations with the updated optical properties. Iterations are performed until the predicted data agrees with the detector readings. The final distribution of optical properties is then displayed as an image.

Non-linear inversion schemes are frequently ill-posed. Model covariance regularisation overcomes the problem of ill-posedness at the expense of limiting the allowed models to a class of model that is compatible with the provided model covariance information. Functional optimisation techniques are highly desirable because they allow the straightforward inclusion of model covariance constraints by including additional terms in the error functional. For calculations performed in this thesis, two additional functional terms have been added to the data misfit functional, penalising both structure and any deviation from the starting model [74, 75]. It is also desirable to use second order derivative optimisation schemes such as the Levenberg-Marquardt method over first order optimisation schemes such as the non-linear conjugate gradient method, since Levenberg-Marquardt calculations treat regularisation terms implicitly. Although Levenberg-Marquardt methods involve computationally time consuming calculations involving both Jacobian and approximate Hessian matrices, the assembly and storage of the Jacobian and approximate Hessian matrix may be bypassed alongside the need to implicitly invert the approximate Hessian. This is achieved using preconditioned conjugate gradient (PCG) solvers to iteratively solve the associated least squares problem. We therefore utilise a variant of the Levenberg-Marquardt optimisation technique, which contains additional regularisation terms to solve the Boltzmann transport equation.

1.1 Scope and Objectives of the Thesis

A main contribution of this thesis is the development of a differentiable discretisation of the Boltzmann transport equation. The discretisation is designed specifically for inverse problems and produces accurate solutions for both optically thick (diffusive) and transparent media. Another contribution is the overall inversion method outlined above. An inversion scheme has been developed for modelling both steady state and time dependent radiation transport.

This thesis describes the reconstruction of the optical properties inside a host medium by proffering investigative calculations. Chapter 2 describes the conventional medical imaging tools used routinely in medical diagnostics and common numerical methods used to describe radiation transport. Chapter 3 describes how the transport of neutral particles may be described according to the one-speed, time-dependent Boltzmann transport equation. The Boltzmann transport equation is discretised in to a six dimensional phase space comprising of spatial ($\mathbf{r}$ = x, y, z), angle ($\mathbf{\Omega} = \theta, \omega$) and time (t) components. A Streamline Upwinded Petrov-Galerkin (SUPG) finite element method is used to spatially discretise the Boltzmann transport equation, a Discontinuous Galerkin (DG) method is used to discretise the time domain, and spherical harmonic basis functions are used to represent the angle of photon travel.

Chapter 4 describes the inverse problem [77, 78, 79] as the reversal of a cause effect sequence, where material properties can be determined from the known dynamics of a physical system. Both first and second order adjoint techniques are considered along with a description of various gradient based optimisation schemes. Chapter 5 provides a detailed investigation into optical imaging using the steady state form of the Boltzmann transport equation, while Chapter 6 provides a series of applications that demonstrate the enhanced imaging achievable using the time dependent form of the Boltzmann transport equation. Chapter 7 is an applications section that models optical imaging through voids using both steady state and time dependent information along with full radiation transport theory and diffusion theory. Chapter 8 provides a summary of our findings and provides recommendations for future work.

Chapter 2

Medical Imaging Techniques

Contents

2.1 Introduction

The most reliable way of diagnosing disease is to look for changes in tissue samples that have been taken from a patient during a biopsy [12]. There is an increasing need for low cost diagnostic tools that are capable of probing tissue samples without the need of invasive surgery. Medical imaging techniques provide such methods by reconstructing the spatial distribution of a selected tissue region from measurements taken outside of the body [80]. Apart from diagnosis, medical imaging also serves to assist with planning and monitoring the treatment of disease [12, 80] and in recent years has advanced insofar that both structural and functional imaging are now possible. Conventional imaging tools used routinely in medical diagnostics include x-ray imaging [46, 80, 81] x-ray computed tomography, (CT) [46, 80, 81, 82, 83] ultrasonic imaging [46, 80, 84, 85, 86], magnetic resonance imaging,(MRI) [46, 80], and nuclear medicine [46, 80, 87, 88, 89].

2.2 X-ray Imaging and X-ray Computed Tomography

X-ray radiography [46, 80, 81] is a transmission based technique in which x-ray photons pass from a source, through an opaque object and are detected, either on a photographic film, or in an ionisation chamber. In biomedicine, contrasts in x-ray imaging occur between different tissue types because of the differential attenuation of x-rays in the body. X-rays are strongly attenuated by bone but readily penetrate soft tissue. Useful information is contained in the transmitted photons that pass straight through the body, before being received by detection apparatus. Scattered photons are also received by the detectors but cause detrimental effects to image quality such as loss of contrast and image artefacts. Radio-opaque contrasting agents -such as iodine or barium based compounds- are frequently administered in soft tissue imaging to enhance image contrast and anti-scatter grids are often used to reduce the amount of unwanted scatter [80].

The x-ray image that is produced is a two dimensional projection (a flat image) of tissue lying between the x-ray source and the detector. Although x-ray imaging can detect and image an abnormality it cannot perceive depth and is therefore limited by its two dimen-

sional projection capabilities. It is also difficult to distinguish different tissue types as often their linear attenuation coefficients are similar. Clinical methods and applications of x-ray radiographic imaging include:

1. The Skeletal System [80] - Since x-rays are strongly attenuated by bone but pass fairly readily through soft tissue, x-ray radiography is particularly useful for imaging the human skeleton. Skeletal imaging is often necessary for analysing the severity of fractures and re-setting broken bones.

2. Vascular Imaging [80] - X-ray imaging is used to study compromised blood flow in the brain and heart as well as the peripheral venous and arterial systems. An iodine based visual contrasting agent is usually used in vascular imaging.

3. Mammography [27, 80] - An x-ray is taken under compression to detect small lesions in breast tissue.

4. The Abdominal System [80] - X-ray scans are used to image the liver, bladder, abdomen and pelvis and to detect disease in the genitourinary tract. Kidney, urethra and bladder (KUB) scans detect the abnormal distribution of gas within the intestines, which is indicative of certain gastro-intestinal conditions and kidney stones. A functional imaging application is Intravenous Pyelography (IVP). IVP uses an iodinated contrasting agent to visualise the filling and emptying of the urinary system. A series of images are acquired at different times after the contrasting agent is administered.

X-ray computed tomography, (CT), [46, 80, 81, 82, 83] is used to image overlapping layers of soft tissue and complex bone structures which are not easily imaged using x-ray radiography. CT images have high spatial resolution and provide better contrast between soft tissue compared to conventional x-ray imaging methods.

During x-ray computed tomography [46, 80] a planar (one dimensional) slice of the body is defined and examined from multiple angles using a pencil or cone like x-ray beam. The x-ray sources and detectors are synchronously rotated, around the patient, encompassing a wide range of viewing angles. The data acquired by the detectors is then passed through a

reconstruction algorithm to produce a two dimensional image. To provide more accurate imaging, a window control is often used to amplify contrast between soft tissue with differing attenuation coefficients. Clinical methods and applications of x-ray computed tomographic imaging include:

1. Cerebral Scans [80, 81, 82] - CT methods have been used to analyse skull fracture, brain damage, stroke, and haemorrhage. Haemorrhage appears on CT scans as an area or increased signal intensity whereas stroke is perceived as an area of reduced signal intensity. CT can also image calcifications, which can be indicative of brain tumours. Iodinated compounds are often used to increase signal intensity during tumour imaging.

2. Pulmonary Disease [80, 82] - CT is used to diagnose both malignancies and diffuse diseases of the lung such as silicosis, fibrosis, cystic fibrosis and emphysema.

3. Abdominal Imaging [80, 81] - CT is used to diagnose abdominal tumours and ulcerations of the liver. Compound fractures in organs such as the pelvis have also been visualised in three dimensions using CT techniques.

4. Angiographic Imaging [80, 82] - High resolution images of blood flow have been accomplished using CT methods.

Both x-ray imaging and x-ray computed tomography use ionising x-radiation. Ionising radiation can cause tissue damage; therefore exposure to such radiation should be kept to a minimum. Radiation dose is of particular concern in paediatric and obstetric radiology.

2.3 Ultrasonic Imaging

Ultrasonic imaging [46, 80, 84, 85, 86] is a non-invasive, portable, and relatively inexpensive diagnostic modality that is capable of real-time functional imaging. Images are constructed from ultrasonic radiation that is backscattered at tissue boundaries and from small structures within tissue. During ultrasonic imaging [46, 80] a short, narrow pulse

(1-5 μs long) of ultrasound is emitted by a transducer, which is placed in close contact with the skin. The ultrasonic pulse propagates as a pressure wave, through tissue at a speed of approximately 1540m/s (the speed of sound). Tissue interfaces, boundaries between tissue and structures within organs often have differing acoustic refractive indices, which causes ultrasound to scatter. The fraction of ultrasound that is backscattered along its original path is received by the original transducer as an echo. The transducer converts the backscattered pulse in to a series of voltages, which are amplified, filtered and digitised instantaneously to produce a moving image.

The depth and location of tissue interfaces are deduced from the time elapsed between the emitted pulse and the received echo, and the propagation velocity of sound in the transmitting medium. Low frequency (1-3MHz) ultrasound penetrates deep into the human body and is therefore used to image deep lying structures such as internal organs. High frequency (5-10MHz) ultrasound has high intrinsic spatial resolution but low penetration depth and is therefore used to image tissue near the surface of the body. Ultrasonic imaging uses non-ionising radiation and is therefore utilised in sensitive regions such as to visualise the eye and the pregnant abdomen. There are some concerns regarding the radiation effect of ultrasound on developing tissue so acoustic energy and exposure time are always kept to a minimum. Clinical applications using ultrasonic radiation are limited because of poor tissue contrast and the ability of gas and bone to impede ultrasound. Clinical applications include:

1. Obstetrics and Gynaecology [80, 85] - Ultrasound is used to monitor foetal health. Parameters such as head size, the condition of the brain ventricles, blood velocity, and spinal condition are commonly measured to check whether a foetus is healthy. Needle guided ultrasonic methods have been used to detect hereditary disease such as Downs Syndrome.

2. Breast Imaging [27, 80, 85] - Ultrasound is commonly used in conjunction with x-ray mammography to diagnose breast cancer. Mammography is used to image tumours and lumps while ultrasound analyses the tissues acoustic properties to determine whether the lump is solid or a fluid filled cyst. Ultrasound is especially useful for younger women whose breast tissue can be relatively opaque to x-rays.

If a needle biopsy is required to test for malignancy ultrasonic imaging can be used to guide a needle into a tumour.

3. Cardiac Illness [80, 85] - Ultrasonic techniques are used to diagnose diseases such as regurgitation, congenital heart disease and cardiac tumours. Doppler ultrasound techniques have been used to measure blood velocity in the arteries and veins in the heart and to detect compromised blood flow. Using contrasting agents in conjunction with ultrasound has allowed the construction of blood perfusion maps.

4. Further applications include intra-abdominal imaging of the liver , kidneys, gallbladder and spleen [80].

2.4 Nuclear Medicine

Most pathological conditions are caused by a change in the biochemistry of tissue. Chemical changes can cause deficiencies in organ function and changes to the physical structure of tissue. Nuclear medicine can detect abnormalities very early in the progression of disease, providing valuable structural and functional imaging capabilities.

Rather than providing an anatomical map, nuclear medicine relies on the ingestion of a radioactive compound (radiopharmaceutical) [87] -via inhalation, direct injection, subcutaneous administration or oral administration- to provide a spatial visualisation of the human body [87]. The radiopharmaceutical turns the system or organ being studied into a (gamma emitting) radioactive source. As the radioactive compound decays the gamma radiation is detected externally using gamma camera equipment.

To deduce the position of the source, a collimator is placed between the patient and the detector so that only gamma radiation with angles close to 90 degrees to the detector plane are recorded. A scintillation crystal then converts the gamma-ray energy that passes through the collimator into light, which, in turn, is converted into an electrical impulse using photomultiplier tubes. The electrical impulse, together with the spatial distribution or the radiopharmaceutical, is analysed by computational apparatus to form an image. Abnormal tissue distribution or an increase or decrease in the rate at which the radiophar-

maceuticals accumulate in particular tissues, organs, or systems are indicative of disease. The health effects associated with ingesting radiation emitters is a major concern.

There are two conventional nuclear imaging methods, both of which depend on the type of radioactive isotope used. Single photon emission computed tomography (SPECT) [80, 88] uses single photon emitters as radioactive sources while positron emission tomography (PET) [80, 89] uses positron emitting radiopharmaceuticals.

2.4.1 SPECT

SPECT [80, 88] scans are performed from various angles using one or more gamma cameras mounted on a computer controlled rotating gantry. The gantry circles the patient at predetermined increments to form a two dimensional image of the tracer distribution. Clinical applications that use SPECT include:

1. Brain Imaging [80] - Regions that show higher uptake in radiopharmaceuticals compared to surrounding tissue often represent tumours. Conversely stroke patients show a lack of blood flow in the effected area of the brain. Blood perfusion in the brain has also been studied. Diseases that cause altered perfusion patterns include epilepsy, cerebral infarction, schizophrenia and dementia.

2. Bone Scanning and Tumour Detection [80, 87] - Whole body scanning has been used to detect bone tumours and soft tissue tumours that cause the deformation and remodelling of bone structure. Soft tissue tumours are characterised by a local uptake of a radiopharmaceutical. Various forms of cancer can also be diagnosed using SPECT, including Hodgkin's disease, lung cancer, non-Hodgkin's lymphoma, malignant melanoma, thyroid cancer, leukaemia, pancreatic tumours and breast and ovarian cancer.

3. Cardiac Imaging [80, 87] - Functional SPECT imaging has been used to measure blood flow in the heart and to detect coronary artery disease. Blockages are visualised as areas of low signal intensity.

4. The Respiratory System [80] - Respiratory diseases can cause disruptions to blood flow, air flow or both. If the pulmonary artery or one of its branches is blocked then radioactivity is absent from that region. Obstructions such as a pulmonary embolism result from abnormal blood flow while abnormal ventilation is indicative of an obstructed airway. Diseases such as bronchitis, asthma and pulmonary edema are characterised by both abnormal blood flow and abnormal ventilation.

5. Liver and the Reticuloendothelial System [80] - Diseases such as cirrhosis of the liver have increased radioactivity in the spleen and bone marrow. Metastatic tumours, cysts, abscesses and haematomas have reduced signal intensity.

6. Renal Imaging [80] - Abnormal functional kinetics are used to detect renal artery stenosis and renal infarction.

2.4.2 PET

The gamma radiation produced during PET imaging [80, 89] is a result of electron-positron annihilation. An annihilation event is only accepted if both photons are detected in coincidence by an array of detectors that surround a patient. Absorbing collimators are therefore not needed. The detection of coincident photons is used to identify the line, along which the original decay occurred, which is subsequently mapped as a two dimensional image. Clinical applications of PET include:

1. Brain Imaging [80] - PET can be used to measure regional cerebral flow and tissue metabolism. High glucose metabolic rates are indicative of malignant gliomas and astrocytomas while low glucose metabolic rates are used in epilepsy diagnosis. A reduction in dopamine synthesis can be indicative of Alzheimer's and Parkinson's disease. Neuroscience uses PET imaging methods to research neural activity.

2. Cardiac Studies [80]- PET imaging can be used to decide whether a heart bypass or heart transplant should be carried out on a patient. Reduced blood flow and metabolism imaging is used to decide whether the heart tissue has died.

3. Tumour Imaging [80] - Malignant cells have higher aerobic glucose metabolism than healthy cells. Metabolic PET imaging is used to diagnose breast, lung, head, and neck cancers.

The absence of absorbing collimators increases the sensitivity of PET imaging by approximately two orders of magnitude compared to SPECT imaging. Compton scattering [90] is particularly problematic during SPECT imaging as it occurs at similar energies to the ingested radiopharmaceuticals and can therefore cause image artefacts or loss of contrast. PET imaging is also advantageous as it has the ability to measure both blood flow and metabolism, using radiopharmaceuticals that generally have shorter lifetimes than SPECT radiopharmaceuticals. However, PET requires a large, expensive cyclotron to produce positron emitting elements.

2.5 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging [46, 80] produces images of the human body using the magnetic resonance signals produced by protons present in water and lipid. In the absence of a magnetic field, protons are randomly orientated. The application of a strong, external magnetic field causes the protons to align in to parallel or anti-parallel configurations. In the lower energy state nuclei align parallel to the direction of the static magnetic field, while in the higher energy state nuclei align antiparallel to the field. The angular momentum of the configured protons causes them to precess around the static magnetic field with a frequency that depends on magnetic field strength. To ensure that precession is coherent a weak radiofrequency field is often applied. The precessing protons are subsequently detected as an induced voltage in a tuned detector coil. The analogue signal is digitised and an inverse two-dimensional Fourier transform [91, 92] is performed to convert the signal into a spatial image. Spatial information is encoded in the magnetic field gradients, the precessional frequencies and the phase of the protons. MRI can be used to detect changes in relaxation times, and water and mineral concentrations, all of which can be indicators of disease. Clinical applications include:

1. Brain Imaging [80] - Changes in iron concentration are indicative of Parkinson's disease, Alzheimer's disease, haemorrhage, and haematomas. Large changes in relaxation times are observed in hydrocephalus and multiple sclerosis.

2. Liver and the Reticuloendothelial System [80] - Hemosiderosis, hemochromatosis and cirrhosis are often diagnosed using MRI. Contrasting agents are used to aid the diagnosis of liver tumours.

3. Musculoskeletal System [80] - Degenerative diseases in the spinal cord and traumatic injury such as disk herniation, disk compression, or epidural haematoma have all been imaged using MRI. Spinal cord tumours and multiple sclerosis can be detected using high resolution MRI imaging.

4. Cardiac System [80] - MRI has been used to detect cardiovascular disease, heart disease, tumours, congenital heart disease and for signs of heart transplant rejection.

The three dimensional capabilities, excellent spatial contrast, high resolution and non-ionising properties make MRI a favourable imaging technique. However MRI has limited application as many molecules can not be imaged with sufficient image resolution or contrast. Contrasting agents are frequently used to increase the contrast between healthy and pathological tissue. The large and expensive equipment used, and the need to keep a patient motionless due to the slow temporal resolution of MRI acquisition also limit its suitability.

2.6 Medical Optical Tomography

The differences in optical properties between healthy and pathological tissue provide contrast when imaging disease [14]. Optical tomography (OT) [2, 3, 4, 6, 12, 46, 93, 94] is a functional imaging tool that uses near infrared radiation, of wavelength 650nm-900nm, to probe biological media. The optical properties of tissue are contained in absorption and scattering coefficients and are exploited to provide qualitative and quantitative images.

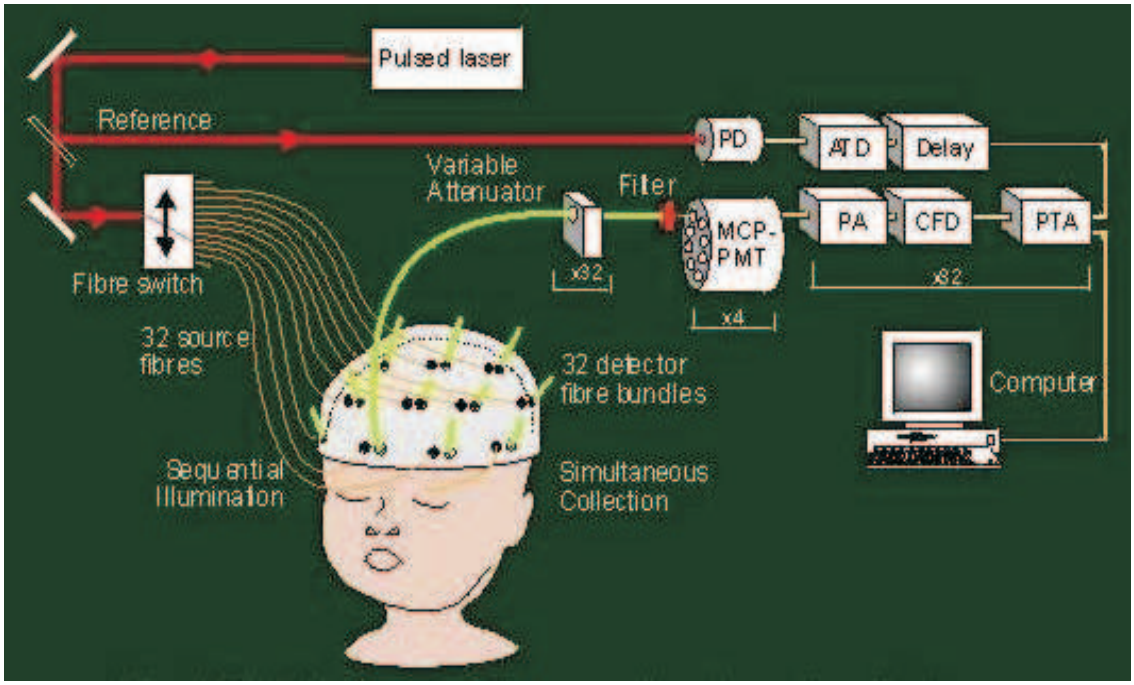


Figure 2.1: A Schematic of an Optical Imaging System [1]

During optical tomographic imaging [2, 6, 7] (figure 2.1), near infrared radiation is guided by fibre optics, which are placed in close contact with the skin. Tissue heterogeneities, boundaries between tissue, and structures within organs often have differing optical refractive indicies which can cause near infrared radiation to scatter. The fraction of scattered photons that emerge from the medium are received by a series of detecting fibres, which are also placed in close contact with the skin. The backscattered pulse is converted into a series of voltages, which are amplified, filtered and digitised instantaneously to produce an image. The data is acquired either as time-varying intensities or as a steady state complex intensity that has measurable amplitude and phase. Clinical applications that use optical radiation include:

1. Brain Imaging [7, 18, 21, 95] - To prevent either brain damage or death by asphyxiation [96, 97, 98], OT can be used to monitor the cerebral blood and tissue oxygenation of new born and preterm infants [19, 20, 22, 99]. OT is also used to detect cerebral haemorrhaging and in the diagnosis of Alzheimer's disease. In Neuroscience OT can be employed to monitor the brain's neural activity during physical or mental exercise [12, 17].

2. Tumour Imaging [12, 46] - Optical imaging could be used to diagnose and monitor skin [100] and breast cancers [28], breast, prostate and brain tumours [28, 29, 30, 31] and brain haematomas.

3. Osteoarthritis, rheumatoid arthritis [23, 24, 101], and diabetes treatments could also benefit from optical monitoring.

Optical tomography is a low cost, non-invasive, functional imaging technique that can be tolerated in large doses without causing tissue damage by ionising or magnetic affects. Near infrared radiation has a low absorption coefficient and can therefore penetrate deeply into tissue [102], providing deep tomographic imaging capabilities. However, scatter dominant light transport [63, 64, 65] is a limiting factor in the development of optical tomography. Each photon undergoes multiple scattering events (deviating considerably from the direct line of sight between the sources and detectors), which causes degradation in image resolution and contrast. The heterogeneous properties of biological tissue [86, 103] also causes difficulties by increasing the dominance of scatter.

A lack of appropriate instrumentation and the absence of sophisticated scattering models, has limited the development of optical tomography as a widely accepted, conventional imaging tool [4]. Predominant scatter makes reconstruction a non-linear problem, prohibiting the use of direct reconstruction methods such as back projection. Instead non-linear, iterative inversion reconstruction schemes are frequently employed [3, 4, 26, 35, 40, 41, 68, 69, 70, 104, 105]. The computational cost of running three dimensional inversion calculations can take hours, limiting the technique to simplistic models and conditions that generalise a more complicated medical system. It is therefore currently desirable to use optical tomography in conjunction with a second imaging modality, like MRI or CT, to monitor changes in critical parameters or to facilitate ongoing diagnosis and research. A method like MRI could provide an initial high resolution map of tissue distribution, which could then be used to guide the computational reconstruction of the low cost, non-invasive optical image.

2.7 Numerical Methods Used to Model Photon Transport

The propagation of light through tissue is governed primarily by absorption and scattering interactions [63, 65]. Biological tissue constitutes a highly scattering medium [106], in which a short pulse of near infrared radiation becomes diffuse over a range of a few millimetres [107, 108]. The propagation of light through biological tissue is fundamentally described according to electromagnetic theory but can be simplified to (radiation) transport theory [32, 109, 110] by ignoring wave phenomena such as polarisation and interference, and particle effects such as inelastic collisions. Two numerical techniques exist that model light propagation in highly scattering media:

1. Stochastic Methods - Stochastic methods, such as Monte Carlo [39, 111, 112, 113, 114], are statistical models that consider the average behaviour of a number of individual particle paths inside a host medium. The path a photon takes is determined by the probability of either a scattering or an absorption event occurring.

2. Deterministic Methods - Deterministic methods rely on solving the Boltzmann transport equation [32, 109, 112] using direct numerical discretisation techniques. Numerical discretisation is performed by dividing the solution domain in to a finite number of regions (elements). In a seven-dimensional phase-space comprising of position $(\mathbf{r} = x, y, z)$, angle $(\mathbf{\Omega} = \theta, \omega)$, energy (E) and time (t) components a number of discretisation techniques exist. Finite difference (FDM) methods [32, 112, 115] result in an orthogonal spatial mesh while control volume (CV) [59, 60, 116] and finite element (FEM) [117, 118, 119, 120] methods are approximated more accurately using unstructured meshes. Angular discretisation [121, 122] is usually achieved using either spherical harmonic (P_N) or discrete ordinate (S_N) approximations [32, 34, 112, 121, 122], while time differencing schemes and multigroup energy schemes are commonly used to discretise the time and energy variables [112].

Deterministic models provide faster solution methods compared to stochastic techniques. However, stochastic techniques provide more accurate results because of the errors associated with numerical discretising a system of deterministic equations.

2.7.1 The Monte Carlo Method

The Monte Carlo method [39, 111, 112, 113, 114, 123] is a stochastic technique in which the life histories of a number of photons are individually tracked from the time of their birth, throughout their various interactions with the host medium, until they are either absorbed into the media or leak out of the medium. Photons that are transmitted into the detector constitute the signal that would be measured during experimental investigations. Since Monte Carlo is a statistical technique, the recorded scattering, transmission and absorption profiles approach true (experimental) values as the number of photons propagated approaches infinity.

During Monte Carlo simulations a model of particular measurement geometry and optical property distribution is sought. Photons are generated at a pre-defined source point and are introduced into the medium one-by-one. To analyse whether a photon is scattered or absorbed, a probability distribution is sought [112]. A random number generator is used to generate both the probability distribution and the direction of photon travel, should a scattering event occur. Once a photon is lost a new photon with a new random number is traced from the initial source point.

Absorption is modelled either by terminating the absorbed photon's path or by introducing a weighting scheme [112, 124]. For models that utilise weighting schemes, the photon packet is split into two parts; a fraction of the photon energy is absorbed, while the rest is scattered with a newly assigned photon weight. Photon life histories are terminated either when the photon path becomes negligible, the photon leaves the domain of interest, or the photon hits the detector. A technique called roulette [112, 124] is often used to terminate the photon's path if the photon weight drops below a specified minimum. Roulette allows a photon one chance in m of surviving with a weight mw otherwise the weight is reduced to zero. The photon is therefore killed in an unbiased fashion without sacrificing energy conservation. If the photon exits the medium, and is measured by a detector, the count rate is increased by the remaining photon weight. If the photon is absorbed into the medium, the position of the absorption is recorded. After sampling many life histories, the average statistical behaviour of the photon particles is found.

The Monte Carlo method is conceptually simple and is particularly useful for modelling complex surfaces (that are not naturally represented by cartesian, cylindrical, or spherical coordinates) and configurations that have little or no symmetry. In principle, Monte Carlo can model the geometry of a problem domain without discretisation error and can incorporate both time dependency and the continuous or discretised treatment of the energy variable. Wave phenomena, such as polarisation and interference, and optical inhomogenities may also be easily incorporated into the Monte Carlo model. However, to obtain adequate statistics, a large number of photon life histories must be tracked, making calculations slow and computationally expensive. Variance reduction techniques [112, 124] can be used to reduce statistical error and therefore increase computation speeds. However, the increased computational speed associated with using variance reduction is still inadequate for many model applications. Despite, computational inefficiencies, Monte Carlo is still the most accurate method for modelling photon transport and is therefore the standard model against which all other numerical methods are benchmarked.

2.7.2 Deterministic Methods

Deterministic methods are principally described according to transport theory [32, 109, 110, 112]. The Boltzmann transport equation can be expressed either in its first order [59, 112] form, its second order [46, 112, 125] form or in its integral [112] form:

1. The first order form of the Boltzmann transport equation [59, 112] is found using deterministic discretisations. The angular variable is usually discretised using discrete ordinates or spherical harmonic discretisation methods [32, 34, 112, 121, 122], while spatial discretisation is achieved using finite differences [32, 112, 115] (e.g. weighted diamond differencing or linear discontinuous methods), finite elements [117, 118, 119, 120] (e.g. Galerkin and characteristic methods) or control volume [59, 60, 116] (e.g. nodal and corner balance methods) based techniques.

2. The most popular second order form of the Boltzmann transport equation is the even parity form [46, 112, 125]. The even parity form is generated by separating the angular flux into its even and odd parity components, thus yielding two coupled

first order differential equations. By discarding all odd parity components, the second order, even parity form of the transport equation is derived. The even parity equation has a variational finite element formulation [118, 126]. Another second order form of the Boltzmann transport equation is the Self-Adjoint Angular Flux (SAAF) equation [127] and is derived using a generalised least squares approach.

3. The integral form of the Boltzmann transport equation [112] involves integrating out any angular dependence. Integral techniques include the collision probability method and the point kernel method. Both methods have limited applicability and are restricted to specific areas or research such as lattice cell calculations and (nuclear) reactor shielding analysis.

This thesis focuses solely on the first order form of the Boltzmann transport equation. Although the second order, even parity, transport equation halves the number of angular unknowns and results in symmetric, positive definite, linear equations [125], it is unable to adequately describe streaming phenomena or materials with large differences in material cross sections [59]. The first order form is therefore a more accurate method for modelling photon transport at tissue interfaces where material properties may vary substantially.

2.8 Deterministic Spatial Discretisation Techniques

Deterministic methods rely on solving the Boltzmann transport equation [32, 109, 112] using direct numerical discretisation techniques. Numerical discretisation is performed by dividing the solution domain in to a finite number of regions (elements). Popular spatial discretisation schemes include, the finite element method [117, 118, 119, 120], the finite difference method [32, 112, 115], and control volume [59, 60, 116] based methods.

2.8.1 The Finite Element Method (FEM)

The finite element method (FEM) [117, 118, 119, 120, 128] is a deterministic technique used to solve partial differential equations posed in complex, and irregular geometries.

The finite element method subdivides a problem domain into a finite number of sub regions called elements, inside which, the solution is approximated using a local, piecewise polynomial basis function. Nodes are typically placed at the corner of each element and the element discretised equations are assembled together to form a global set of linear equations. Provided a set of pre-defined boundary conditions are satisfied [32, 46, 59, 112, 125], a unique trial solution may be found to the system of algebraic equations. The trial solution may then be placed into the original set of differential equations and solved using either weighted residual [117, 118] or variational techniques [117, 118, 126].

The Weighted Residual Method

The weighted residual method [117, 118] utilises the weak (integral) form of the Boltzmann transport equation. An error (defined as the residual) exists between the trial solution and the true solution once the trial solution is replaced into the original Boltzmann transport equation. The weak form is obtained by multiplying the Boltzmann transport equation by an arbitrary weight function and integrating the entire expression over the solution domain. An exact solution is found when the residual is zero at all points along the domain. The most popular weighted residual methods are Point Collocation, Sub Domain Collocation, Least Squares and Galerkin methods [58, 59, 117, 118, 129]. The only difference between each of these weighted residual methods is the value that is chosen to represent the weight. If the weight function is chosen to be the same value as the basis functions, then the weighted residual approach is a Petrov-Galerkin formulation [58, 59, 118, 129].

The Variational Method

The variational principle [46, 117, 118, 125, 126] is a means to find a function with specified values at end points such that the integral between the end points is stationary. A trial function is approximated and is varied until a minimum potential energy is found. The trial functional should satisfy any specified continuity or boundary condition. Once

a minimum is found, the solution is approximated along each element and summed along all elements to give an approximate solution across the domain. Variational methods yield a functional with lower order derivatives than differential methods and can be used to prove the existence of a solution. A variational form can be derived from a differential equation using an Euler-Lagrange formulation, bi-linear functional techniques or least-squares methods [46, 117, 118, 125, 126].

2.8.2 The Finite Difference Method (FDM)

The finite difference method [32, 112, 115, 130] utilises a uniformly spaced, orthogonal (Cartesian) grid to solve a partial differential equation. At each node the partial differential equation is approximated using an algebraic expression, which references adjacent nodes. The system is then solved to find a unique solution.

Models that utilise finite element techniques are considerably more complex than finite difference methods. The use of structured, orthogonal grids mean that the range of problems that can be solved accurately using finite difference methods is limited. For complex or irregular geometry, finite element methods model the problem domain more accurately than finite difference methods and therefore yield more accurate solutions.

2.8.3 The Control Volume Method (CV)

The control volume method [59, 60, 116] involves control volume integration over a finite volume mesh. Control volume techniques are similar to finite element methods in that unstructured meshes are used to discretise the spatial domain. While finite difference and finite element methods represent particle flow at a series of nodal points, the finite (control) volume method defines the boundaries (or faces) of each control volume, midway between adjacent nodes. Each node is therefore surrounded by a control volume cell. To calculate the flux at a given control volume an approximate distribution of properties between nodal points is used. Approximations are made by utilising either local piecewise polynomial (basis) functions [118] or algebraic expressions. The total flux is

then approximated by calculating the convective fluxes through the faces of the control volume. In other words a balance equation is solved across each control volume element.

Chapter 3

Photon Propagation in Optical Media

Contents

3.1 Introduction

To investigate photon migration time and perform inverse scattering calculations that compensate for the degradation in image resolution caused by scatter, it is necessary to study the propagation of light in biological tissue as a function of different optical properties [63, 64]. Interaction between light and tissue is fundamentally described using analytic theory [65, 109, 110, 131]. Analytic theory describes the propagation of photons according to the wave equation of classical electrodynamics (i.e. according to Maxwell's equations), thereby taking into account all coherent optical effects such as interference, diffraction and polarisation [90]. However, due to the mathematical complexity of the analytic model, the inhomogeneity of tissue and the lack of detailed information on the dielectric properties of biological tissue, the analytic model is rarely used to model photon migration in biological media. Instead, transport theory [32, 131] is used as an approximation to analytic theory. Transport theory describes the migration of neutral particles according to the Boltzmann transport equation [32, 112]. Transport theory lacks the rigour of analytic theory by ignoring the wave properties of light. As such, the flow of neutral particles through a medium of randomly scattering and absorbing particles is considered.

3.2 The Optical Parameters of Tissue

The propagation of light in tissue is governed by the transport of individual photons, which are affected by the material properties - namely the absorption and scattering characteristics - of that tissue [64] . Tissue properties are typically characterised by an absorption coefficient, a scattering coefficient, a reduced scattering coefficient and a scattering phase function that represents angular distribution [32, 65, 109, 112, 131].

3.2.1 The Absorption Coefficient

During absorption a photon is removed from the incident beam and is absorbed by a particle in the medium. The removal of a photon reduces the intensity of the incident beam.

The frequency of absorption events is described by the absorption coefficient, $\mu_a(\mathbf{r}, E)$, and is expressed as the product of the atomic number or particle density contained in the material, $\mathcal{A}(\mathbf{r})$, and the microscopic absorption cross-section, $\sigma_a(E)$:

$$\mu_a(\mathbf{r}, E) = \mathcal{A}(\mathbf{r})\sigma_a(E), \tag{3.1}$$

where $\mathbf{r}$ represents a position variable and E represents the energy variable. The reciprocal, $1/\mu_a(\mathbf{r}, E)$ is called the absorption path length and represents the mean free path a photon travels between successive absorption events.

3.2.2 The Scattering Coefficient

During scattering a photon collides with a particle in the medium and scatters into a different direction, resulting in a decrease in the initial beam intensity. The frequency of scattering events is described by a scattering coefficient, $\mu_s(\mathbf{r}, E)$ and is expressed in terms of atomic number or particle density, $\mathcal{A}(\mathbf{r})$, and microscopic scattering cross section, $\sigma_s(E)$:

$$\mu_s(\mathbf{r}, E) = \mathcal{A}(\mathbf{r})\sigma_s(E). \tag{3.2}$$

The reciprocal, $1/\mu_s(\mathbf{r}, E)$, is called the scattering path length and equals the average distance a photon travels between consecutive scattering events.

3.2.3 The Total Attenuation Coefficient

The total attenuation, $\mu_t(\mathbf{r}, E)$, of an incident photon beam is expressed in terms of the total absorption and scattering contributions.

$$\mu_t(\mathbf{r}, E) = \mu_a(\mathbf{r}, E) + \mu_s(\mathbf{r}, E). \tag{3.3}$$

3.2.4 The Differential Scattering Coefficient

The differential scattering cross-section accounts for the angular dependence of the photons emerging from a collision. The differential scattering cross section, $\mu_s(\mathbf{r}, E^{'} \rightarrow E, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega})$, is best described as the probability that a photon moving initially in a direction $\mathbf{\Omega}^{'}$ is scattered into a new direction $\mathbf{\Omega}$ [64].

$$\mu_s(\mathbf{r}) = \int_0^\infty dE \int_\Omega \mu_s(\mathbf{r}, E^{'} \rightarrow E, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega}) d\mathbf{\Omega}. \tag{3.4}$$

The normalised form of the differential scattering cross-section is called the differential phase function, which describes the angular distribution of a single scattering event.

$$f(E^{'} \rightarrow E, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega}) = \frac{\mu_s(E^{'} \rightarrow E, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega})}{\mu_s(E)}. \tag{3.5}$$

The differential phase function has the normalisation condition.

$$\int_0^\infty dE \int_\Omega f(E^{'} \rightarrow E, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega}) d\mathbf{\Omega} = 1. \tag{3.6}$$

For situations in which the scattering phase function is not applicable or not known, the anisotropy factor, g, is instead used to describe the average cosine of the scattering angle, θ. Assuming scattering is independent of the incident direction of travel, $\mathbf{\Omega}^{'}$, and the scattered photon direction, $\mathbf{\Omega}$, and g depends on the absolute value of the angle, θ, spanned over the incident and scattered particle directions, the scattering phase function becomes:

$$f(E^{'} \rightarrow E, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega}) = f(E^{'} \rightarrow E, \mathbf{\Omega}^{'} \cdot \mathbf{\Omega}) = f(E^{'} \rightarrow E, cos\theta) = f(E^{'} \rightarrow E, \mu_0). \tag{3.7}$$

Furthermore, the mean cosine of the scattering angle is subsequently described according to the anisotropy factor, g.

$$g = \frac{\int_0^\infty dE \int_\Omega (\mathbf{\Omega}' \cdot \mathbf{\Omega}) f(E' \rightarrow E, \mathbf{\Omega}' \cdot \mathbf{\Omega}) d\mathbf{\Omega}}{\int_0^\infty dE \int_\Omega f(E' \rightarrow E, \mathbf{\Omega}' \cdot \mathbf{\Omega}) d\mathbf{\Omega}}. \tag{3.8}$$

The anisotropy factor can be assigned numbers between -1 and 1 to describe highly backward ($g \approx -1$), isotropic ($g = 0$) and forward scattering ($g \approx 1$) [64]. In biological media scattering is typically highly anisotropic and forward directed [64].

3.2.5 The Reduced Scattering Coefficient

The reduced scattering coefficient is equivalent to an isotropic scattering coefficient. The reduced scattering coefficient describes light propagation in diffusely scattering media (such as biological tissue) as a cumulative effect of a number of forward scattering events. Combining scattering anisotropy, g with the scattering coefficient, $\mu_s(\mathbf{r}, E)$, yields the reduced scattering coefficient, $\mu_s'(\mathbf{r}, E)$:

$$\mu_s'(\mathbf{r}, E) = \mathcal{A}(\mathbf{r})\sigma_s'(E) = \mu_s(\mathbf{r}, E)\,(1 - g)\,. \tag{3.9}$$

The mean path length travelled by a collimated beam of light before it becomes diffuse is equal to, $1/\mu_s'(\mathbf{r}, E)$.

3.2.6 Optical Transmission Characteristics of Biological Media

The absorption characteristics of biological tissue [137] are mainly influenced by the contents of protein, haemoglobin and water [7, 63, 64, 102, 138]. Blood absorption is strong for visible light until wavelengths reach 600nm at which point absorption decreases dramatically. Protein absorption decreases as wavelength increases and water is found to be a high absorber for deep ultraviolet wavelengths (150-400nm) and again for wavelengths greater than 1500nm. The wavelengths that best describe optical transmission through tissue are therefore contained in the infrared and near infrared range (600-1500nm) [4, 38, 64, 100, 139, 140, 141, 142, 143, 144, 145, 132, 133, 134, 135, 136, 146].

Tissue Type	Sample	$\lambda[nm]$	$\mu_a[mm^{-1}]$	$\mu_s^{'}[mm^{-1}]$	Reference
Adult brain	in vitro	700-900	0.1-0.2	2-5	[132]
Adult grey matter	in vitro	633	0.263	0.722	[64]
Adult white matter	in vitro	800	0.006	8.50	[133]
Adult white matter	in vivo	849	0.013	0.98	[134]
Adult white matter	in vitro	650-900	0.02-0.03	8-10	[133]
Adult white matter	in vitro	633	0.158	0.204	[64]
Adult grey matter	in vitro	800	0.035	2.20	[133]
Adult grey matter	in vivo	811	0.018-0.019	0.48-0.74	[134]
Adult grey matter	in vivo	849	0.018-0.019	0.45-0.74	[134]
Adult grey matter	in vitro	650-900	0.04-0.06	1.9-2.2	[133]
Neonatal Brain	in vitro	800	0.0215	0.748	[133]
Neonatal Brain	in vitro	800	0.0373	0.673	[133]
Neonatal white matter	in vitro	800	0.0373	0.659	[133]
Neonatal white matter	in vitro	650-900	0.04-0.07	0.5-1.2	[133]
Neonatal Grey Matter	in vitro	800	0.0460	0.529	[133]
Neonatal grey matter	in vitro	650-900	0.04-0.08	0.4-0.9	[133]
Cerebrospinal Fluid		800	0.0022	≈ 0	[135]
Cerebrospinal fluid		650-900	0.001	0.01	[4]
Adult skull	in vivo	849	0.022	0.91	[134]
Pig skull	in vitro	800	0.025	1.8	[136]
Pig skull	in vitro	650-950	0.04-0.05	2.63-1.32	[136]
Skin: epidermis	in vitro	400-800	2.4-0.02	3.2-2.1	[100]
Skin: dermis	in vitro	800	0.013	2	[38]
Skin: dermis	in vitro	633-900	0.033-0.013	2.73-1.63	[38]
Skin: subdermis	in vitro	800	0.008	1.2	[38]
Skin: subdermis	in vitro	633-900	0.013-0.009	1.26-1.08	[38]
Medulloblastoma (tumour)	in vivo	849	0.008	0.66	[134]
Astrocytoma (tumour)	in vivo	849	0.019	0.59	[134]

Table 3.1: Optical properties of various brain tissue types taken at NIR wavelengths

Besides protein, haemoglobin and water, biological tissue is composed of a number of chromophores including melanin, cytochrome C oxidase, and myglobin. These chromophores contribute little attenuation at infrared and near infrared wavelengths and are therefore largely ignored in cross-section calculations [7].

The wavelength range between 600-1500nm is often referred to as the optical window. During experimental optical tomography [7] the useful range of the optical window is typically 650-900nm. Wavelengths smaller than 650nm exhibit increased absorption, which contributes to beam attenuation, while there is a rapid decrease in the quantum efficiency of practical detectors at wavelengths greater than 900nm. The most significant beam attenuation inside the optical window is caused by light scattering against cells and organelles.

Table 3.1 summarises the absorption and scattering coefficients of human brain tissue types together with the corresponding references. Note the lower scattering coefficients associated with the neonatal human brain compared to the adult human brain. Smaller brain tissue scattering coefficients together with a smaller head size, a thinner skull and a thinner layer of CSF fluid -a thin layer of virtually transparent fluid that cushions the human brain- suggests that light is far more likely to penetrate deeply into the brain of a neonate as compared to an adult. As such, experimental optical tomography has focussed primarily on neonatal brain imaging as a simplified model [7].

3.3 The Boltzmann Transport Equation

The Boltzmann transport equation describes the transport of neutral particles (such as photons and neutrons) using kinetic theory [32]. The transport of optical radiation is therefore analysed in the same manner as neutron transport phenomena [112]. The Boltzmann transport equation models the distribution of particles inside a host medium while taking into account the motion of these particles and their interaction with the host medium [32, 34, 112]. For optical tomography applications the time dependent Boltzmann transport equation is used to describe the angular flux, $\psi(\mathbf{r}, \mathbf{\Omega}, E, t)$, that results from an ex-

ternal source of photons, $s(\mathbf{r}, \mathbf{\Omega}, E, t)$, incident on a host medium. The time dependent Boltzmann transport equation is written:

$$\frac{1}{v}\frac{\partial\psi(\mathbf{r}, \mathbf{\Omega}, E, t)}{\partial t} + \mathbf{\Omega} \cdot \nabla\psi(\mathbf{r}, \mathbf{\Omega}, E, t) + \mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}, E, t) = s(\mathbf{r}, \mathbf{\Omega}, E, t), \qquad (3.10)$$

where $\mathbf{r}$ is the position vector ($\mathbf{r} = x, y, z$) of a particle or photon moving along direction, $\mathbf{\Omega} = (\theta, \omega)$, at time, t, and with energy E. The system has characteristic velocity, v. $\psi(\mathbf{r}, \mathbf{\Omega}, E, t)$ represents the particle angular flux and $s(\mathbf{r}, \mathbf{\Omega}, E, t)$ represents an external source term. The scattering/removal operator, $\mathcal{H}$ is defined:

$$\begin{aligned} \mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}, E, t) &= (\mu_a(\mathbf{r}, E) + \mu_s(\mathbf{r}, E))\psi(\mathbf{r}, \mathbf{\Omega}, E, t) \\ &- \int_0^\infty dE' \int_\Omega \mu_s(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega}, E' \rightarrow E)\psi(\mathbf{r}, \mathbf{\Omega}', E', t)d\mathbf{\Omega}', \end{aligned} \qquad (3.11)$$

where $\mu_a(\mathbf{r}, E)$ and $\mu_s(\mathbf{r}, E)$ represent position dependent absorption and scattering coefficients respectively, and $\mu_s(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega}, E' \rightarrow E)$ represents the differential macroscopic cross-section. The Boltzmann transport equation is a statement of particle conservation [46, 125] . The time rate of change of the number of particles is a balance relation between the various mechanisms through which particles can be gained or lost from the volume element considered. The external source term on the right hand side of equation 3.10 and the second scattering term on the right hand side of equation 3.11 both represent particles gained by the volume element. Each of the other terms represent loss mechanisms.

The complexity of the Boltzmann transport equation as having seven independent variables of position ($\mathbf{r} = x, y, z$), angle ($\mathbf{\Omega}$ = θ, ω), energy, (E), and time, (t), either limits its utility to extremely simple situations (e.g. 1 dimensional problems) or requires the use of numerical approximation methods [34, 46] to generate a solution. To simplify the analysis of the discretisation scheme the particle distribution is assumed to be independent of energy. This reduces the distribution to a one-group, one-speed approximation.

$$\frac{1}{v}\frac{\partial\psi(\mathbf{r},\mathbf{\Omega},t)}{\partial t} + \mathbf{\Omega}\cdot\nabla\psi(\mathbf{r},\mathbf{\Omega},t) + \mathcal{H}\psi(\mathbf{r},\mathbf{\Omega},t) = s(\mathbf{r},\mathbf{\Omega},t). \tag{3.12}$$

To complete the mathematical description of the Boltzmann transport equation, boundary conditions must be defined for a particular problem [32, 34, 109, 112]. Since the Boltzmann transport equation is a first order differential equation in space and time, boundary conditions are required in both space and time variables. There are a number of different types of boundary conditions [32, 109, 112]. Each of the models presented in this thesis have vacuum boundary conditions:

$$\psi(\mathbf{r}_s,\mathbf{\Omega},t) = 0, \quad when \quad \mathbf{n}\cdot\mathbf{\Omega} < 0, \tag{3.13}$$

in which, $\mathbf{n}$ represents the normal to the boundary. In other words, any photon that escapes through the bounding surface is not able to re-enter the medium through another part of the surface.

3.4 Discretisation of the Boltzmann Transport Equation

A discretisation scheme is utilised to reduce the Boltzmann transport equation in to a coupled system of multi-dimensional differential-difference equations. A Streamline Upwind Petrov-Galerkin (SUPG) finite element method is used to spatially discretise the Boltzmann transport equation and spherical harmonic basis functions are used to represent the angle of photon travel. The SUPG method utilises a weighted residual (Petrov Galerkin) method in order to obtain the weak form of the systems equations. An additional contribution to the convection term is also added to the Petrov Galerkin weighting. The additional term induces further diffusion in the streamline direction in order to suppress spurious Gibbs oscillations. SUPG methods are accurate and robust across all radiation transport regimes from optically thick to optically thin media and form a robust coupling between time and space dimensions. The methods apply a finite element treatment for the internal domain and a Riemann approach on the boundary of the domain. This significantly

simplifies the numerical application of the scheme by circumventing the evaluation of complex half range angular integrals on the domain boundary.

3.4.1 The Angular Discretisation of the Angular Flux

Angular discretisation is achieved using a Petrov-Galerkin approach [58, 59, 119, 120]. The Petrov-Galerkin method is applied by multiplying equation 3.12 by a series of angularly dependent finite element weight functions, $M_\rho(\mathbf{\Omega})$, and integrating the subsequent expression over the solution domain. Additionally a finite element approximation $\mathbf{\Psi}(\mathbf{r},t)$, replaces the continuous angular flux, $\psi(\mathbf{r},\mathbf{\Omega},t)$ and a finite element approximation, $\mathbf{S}(\mathbf{r},t)$, approximates the source function, $s(\mathbf{r},\mathbf{\Omega},t)$. The equations become:

$$\int_{\mathbf{\Omega}} M_\rho(\mathbf{\Omega}) \left[\frac{1}{v}\frac{\partial\psi(\mathbf{r},\mathbf{\Omega},t)}{\partial t} + \mathbf{\Omega}\cdot\nabla\psi(\mathbf{r},\mathbf{\Omega},t) + \mathcal{H}\psi(\mathbf{r},\mathbf{\Omega},t) - s(\mathbf{r},\mathbf{\Omega},t) \right] d\mathbf{\Omega} = 0. \tag{3.14}$$

The approximation $\mathbf{\Psi}(\mathbf{r},t)$ for the angular flux, $\psi(\mathbf{r},\mathbf{\Omega},t)$, is described using the finite element basis function, $M_\nu(\mathbf{\Omega})$:

$$\psi(\mathbf{r},\mathbf{\Omega},t) \approx \sum_{\nu=1}^{\mathcal{M}} M_\nu(\mathbf{\Omega})\Psi_\nu(\mathbf{r},t), \tag{3.15}$$

where $\mathbf{\Psi}(\mathbf{r},t) = (\Psi_1(\mathbf{r},t)\ \Psi_2(\mathbf{r},t)\ ...\ \Psi_{\mathcal{M}}(\mathbf{r},t))^T$ represents a vector of $\mathcal{M}$ angular moments and $M_\nu(\mathbf{\Omega})$ represents $\mathcal{M}$ angular basis functions. The approximation, $\mathbf{S}(\mathbf{r},t)$, for the source function $s(\mathbf{r},\mathbf{\Omega},t)$, is also expanded using a discrete series of basis functions:

$$s(\mathbf{r},\mathbf{\Omega},t) \approx \sum_{\nu=1}^{\mathcal{M}} M_\nu(\mathbf{\Omega})S_\nu(\mathbf{r},t), \tag{3.16}$$

where $\mathbf{S}(\mathbf{r},t) = (S_1(\mathbf{r},t)\ S_2(\mathbf{r},t)\ ...\ S_{\mathcal{M}}(\mathbf{r},t))^T$. The angular basis functions are usually expanded using either discrete ordinate, (S_N), or spherical harmonic, (P_N), approximation techniques [32, 34, 109, 112].

The Expansion of the Angular Flux Using Discrete Ordinates

The discrete ordinates, (S_n), method approximates the Boltzmann transport equation by truncating the expression for the angular flux, $\psi(\mathbf{r}, \mathbf{\Omega}, t)$, so that particle travel only occurs in a finite number of directions. The direction of particle travel is found by placing a unit sphere over each spatial point and partitioning that sphere into a finite number, ($\mathcal{M}(\mathcal{M}+2)$), of patches. The angular flux is then described according to a quadrature scheme, which is used to estimate the integral over the angular flux variable.

$$\int_{\Omega} \psi(\mathbf{r}, \mathbf{\Omega}, t) d\Omega = \sum_{\nu=1}^{\mathcal{M}} w_{\nu}(\mathbf{\Omega}) \Psi_{\nu}(\mathbf{r}, t), \tag{3.17}$$

in which, $w_{\nu}(\mathbf{\Omega})$ represents the quadrature weights associated with each moment, ν. Since the quadrature scheme is arbitrary, it can be chosen so that the expected angular flux dependence may be integrated. For example if, the angular dependence is highly forward peaked, the quadrature can be selected so that points are concentrated in the appropriate range. If no prior knowledge of the angular distribution is known, a Gauss-Legendre scheme is often used. For three dimensional problems the angular flux, $\psi(\mathbf{r}, \mathbf{\Omega}, t)$, must be determined over all eight octants of the unit sphere swept out by $\mathbf{\Omega}$. In two dimensional geometries the mirror symmetry of $\psi(\mathbf{r}, \mathbf{\Omega}, t)$ about the plane formed by two orthogonal co-ordinates reduces the number of octants over which the angular dependence must be determined to four [32, 109, 112].

The Expansion of the Angular Flux Using Spherical Harmonics

The spherical harmonics (P_N) method approximates the Boltzmann transport equation by expanding the angular flux as an infinite set of spherical harmonics [32, 34, 109, 112]. For an arbitrary $\mathcal{M}$, angular flux may be approximated by truncating the expansion to retain only the harmonic functions of order $\mathcal{M}$ and below. The spherical harmonic expansion is unbounded and the asymptotic limit (as $\mathcal{M} \rightarrow \infty$) converges to the exact solution of the Boltzmann transport equation. For planar problems, the P_N equations are of order $(\mathcal{M}+1)$, while for multi-dimensional problems the P_N equations grow to order $(\mathcal{M}+1)^2$.

Multi Dimensional Spherical Harmonics

Expanding angular flux as an infinite set of spherical harmonics yields the series [59]:

$$\Psi_{l,m}^{\mathcal{R},\mathcal{C}}(\mathbf{r},t) = \int_{\Omega} \psi(\mathbf{r},\mathbf{\Omega},t) Y_{l,m}^{\mathcal{R},\mathcal{C}}(\mathbf{\Omega}) d\mathbf{\Omega}. \tag{3.18}$$

The superscript $\mathcal{R}$ is used to denote a real orthonormal series, while the superscript $\mathcal{C}$ is used to define the series in complex terms. The spherical harmonic expansion results in a sum of a cosine, $Y_{l,m}^{c}(\mathbf{\Omega})$, and sine series, $Y_{l,m}^{s}(\mathbf{\Omega})$.

$$Y_{l,m}^{\mathcal{C}}(\mathbf{\Omega}) = Y_{l,m}^{c}(\mathbf{\Omega}) + iY_{l,m}^{s}(\mathbf{\Omega}). \tag{3.19}$$

Furthermore, the real, orthonormal $Y_{l,m}^{c}(\mathbf{\Omega})$, and $Y_{l,m}^{s}(\mathbf{\Omega})$ functions are each defined as:

$$Y_{l,m}^{\mathcal{R}}(\mathbf{\Omega}) = \begin{cases} Y_{l,m}^{c}(\mathbf{\Omega}) = C_{l,m}^{1/2}\, P_{l,m}(\mu)\, cos(m\omega), & 0 \leq m \leq l \\ Y_{l,m}^{s}(\mathbf{\Omega}) = C_{l,m}^{1/2}\, P_{l,m}(\mu)\, sin(m\omega), & 0 \leq m \leq l \end{cases}$$

where μ is the cosine of the polar angle in spherical polar co-ordinates, $C_{l,m} = \dfrac{(2l+1)(l-m)!}{(l+m)!}$ and $P_{l,m}(\mu)$ are associated Legendre polynomials [112]. In complex terms, equation 3.19 is defined:

$$Y_{l,m}^{\mathcal{C}}(\mathbf{\Omega}) \;\; = \;\; C_{l,m}^{1/2}\, P_{l,m}(\mu)\, exp(im\omega), \tag{3.20}$$

where $-l \leq m \leq l$. The parity of the spherical harmonic basis functions is dictated by l:

$$Y_{l,m}^{\mathcal{R},\mathcal{C}}(-\mathbf{\Omega}) = (-1)^{l} Y_{l,m}^{\mathcal{R},\mathcal{C}}(\mathbf{\Omega}). \tag{3.21}$$

Odd parity moments are associated with odd values of l, while even parity moments are

associated with even values of l. Expanding equation 3.18 in terms odhtf real, orthonormal sphereical harmonics yields an infinite Fourier-Laplace series:

$$\psi(\mathbf{r},\mathbf{\Omega},t) = \sum_{l=0}^{\infty}\left[\sum_{m=0}^{l}\Psi^{c}_{l,m}(\mathbf{r},t)Y^{c}_{l,m}(\mathbf{\Omega}) + \sum_{m=1}^{l}\Psi^{s}_{l,m}(\mathbf{r},t)Y^{s}_{l,m}(\mathbf{\Omega})\right], \tag{3.22}$$

where the two expansion coefficients $\Psi^{c}_{l,m}(\mathbf{r},t)$ and $\Psi^{s}_{l,m}(\mathbf{r},t)$ each represent an angular variation in the angular flux:

$$\begin{aligned}
\Psi^{c}_{l,m}(\mathbf{r},t) &= \int_{\Omega}\psi(\mathbf{r},\mathbf{\Omega},t)Y^{c}_{l,m}(\mathbf{\Omega})d\mathbf{\Omega}, \;\; 0 \leq m \leq l \\
\Psi^{s}_{l,m}(\mathbf{r},t) &= \int_{\Omega}\psi(\mathbf{r},\mathbf{\Omega},t)Y^{s}_{l,m}(\mathbf{\Omega})d\mathbf{\Omega}, \;\; 0 \leq m \leq l
\end{aligned} \tag{3.23}$$

Spherical harmonics obey the following orthonormal conditions:

$$\begin{aligned}
\int_{\Omega}Y^{s}_{l,m}(\mathbf{\Omega})Y^{s}_{l',m'}(\mathbf{\Omega})d\mathbf{\Omega} &= \delta_{l,l'}\delta_{m,m'}, \\
\int_{\Omega}Y^{c}_{l,m}(\mathbf{\Omega})Y^{c}_{l',m'}(\mathbf{\Omega})d\mathbf{\Omega} &= \delta_{l,l'}\delta_{m,m'}, \\
\int_{\Omega}Y^{c}_{l,m}(\mathbf{\Omega})Y^{s}_{l',m'}(\mathbf{\Omega})d\mathbf{\Omega} &= 0.
\end{aligned} \tag{3.24}$$

The angular Jacobian matrices projected onto a set of spherical harmonic basis functions are finally defined:

$$\mathbf{A}_{k(ll'\ mm')} = \int_{\Omega}Y^{\mathcal{R}}_{l',m'}(\mathbf{\Omega})\ \Omega_k\ Y^{\mathcal{R}}_{l,m}(\mathbf{\Omega})d\mathbf{\Omega}, \tag{3.25}$$

where $k = x, y$, and z components respectively. $\mathbf{A}_k$ is of size $\mathcal{M}\times\mathcal{M}$, where $l = 0,1,...,L$, $l^{'} = 0,1,...,L$, $m = -l,...,l$, and $m^{'} = -l,...,l$. The analytic form of the

angular Jacobian mass matrices, created using multi-dimensional spherical harmonics, is described in [59].

3.4.2 The Expansion of the Scattering Contributions

For non-multiplying systems [112] there are two contributions to emission density that represent particles gained by a given volume. These are contributions due to external sources, $s(\mathbf{r}, \mathbf{\Omega}, t)$, and contributions due to scattered particles, $q_s(\mathbf{r}, \mathbf{\Omega}, t)$. The latter, scattering contributions can be defined according to:

$$q_s(\mathbf{r}, \mathbf{\Omega}, t) = \int_{\Omega} \mu_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}^{'}) \psi(\mathbf{r}, \mathbf{\Omega}^{'}, t) d\mathbf{\Omega}^{'}. \tag{3.26}$$

Scattering contributions are also contained in the total cross-section:

$$\mu_t(\mathbf{r}) = \mu_a(\mathbf{r}) + \mu_s(\mathbf{r}). \tag{3.27}$$

For practical applications it is often assumed that scattering collisions have rotational symmetry. Therefore, the differential scattering cross-section, $\mu_s(\mathbf{r}, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega})$, is not dependent on the incident and scattered particle directions, but instead on the change in particle direction (equation 3.7). Consequently, $\mu_s(\mathbf{r}, \mu_0)$, may be expanded as a series of orthogonal Legendre polynomials [112]:

$$\mu_s(\mathbf{r}, \mu_0) = \sum_{l=0}^{\infty} \frac{(2l+1)}{4\pi} \mu_{sl}(\mathbf{r}) P_l(\mu_0), \tag{3.28}$$

where $P_l(\mu_o)$ is a Legendre Polynomial of order l. The scattering moments, $\mu_{sl}(\mathbf{r})$, are derived from the orthogonality of Legendre polynomials [91, 92, 112] according to:

$$\mu_{sl}(\mathbf{r}) = 2\pi \int_{-1}^{1} \mu_s(\mathbf{r}, \mu_0)\ P_l(\mu_0) d\mu_0. \tag{3.29}$$

The scattering coefficient, $\mu_{sl}(\mathbf{r})$, has the normalisation condition:

$$\mu_{s0}(\mathbf{r}) = 2\pi \int_{-1}^{1} \mu_s(\mathbf{r}, \mu_0) d\mu_0 = 1. \tag{3.30}$$

Replacing equation 3.28 into 3.26 yields:

$$q_s(\mathbf{r}, \mathbf{\Omega}, t) = \sum_{l=0}^{\infty} \frac{(2l+1)}{4\pi} \mu_{sl}(\mathbf{r}) \int_{\Omega} P_l(\mu_0) \psi(\mathbf{r}, \mathbf{\Omega}', t) d\mathbf{\Omega}'. \tag{3.31}$$

Using Legendre addition theorem [91, 112], the integral defined in equation 3.31 becomes:

$$P_l(\mu_0) = P_l(\mathbf{\Omega} \cdot \mathbf{\Omega}') = \frac{4\pi}{(2l+1)} \sum_{m=-l}^{l} Y^*_{l,m}(\mathbf{\Omega}) Y_{l,m}(\mathbf{\Omega}'). \tag{3.32}$$

The asterisk denotes a complex conjugate. Combining equations 3.31 and 3.32 yields:

$$q_s(\mathbf{r}, \mathbf{\Omega}, t) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y^*_{l,m}(\mathbf{\Omega}) \mu_{sl}(\mathbf{r}) \Psi_{l,m}(\mathbf{r}, t). \tag{3.33}$$

in which, $\Psi_{l,m}(\mathbf{r}, t)$ represents the angular variation in the angular flux (see equation 3.23) and results from expanding the angular flux as a spherical harmonic series (according to equations 3.18, 3.19, 3.20, and 3.21). Applying the Galerkin method [119, 120] to the entire system of equations yields:

$$\begin{aligned} q_s(\mathbf{r}, \mathbf{\Omega}, , t) &= \sum_{\nu=1}^{\mathcal{M}} \Psi_\nu(\mathbf{r}, t) \Big[\sum_{l=0}^{\infty} \mu_{sl}(\mathbf{r}) \alpha^c_{\rho,l,0} \alpha^c_{\nu,l,0} \\ &+ 2 \sum_{l=1}^{\infty} \mu_{sl}(\mathbf{r}) \sum_{m=1}^{l} \alpha^c_{\rho,l,m} \alpha^c_{\nu,l,m} + \alpha^s_{\rho,l,m} \alpha^s_{\nu,l,m} \Big]. \end{aligned} \tag{3.34}$$

The bracketted terms are collectively referred to as the scattering matrix. The angular

discretisation of the scattering/removal operator, $\mathcal{H}$, thus yields the diagonal matrix (for spherical harmonics):

$$\begin{aligned}\mathbf{H}_{\rho\nu}(\mathbf{r}) &= \int_{\mathbf{\Omega}} (\mu_a(\mathbf{r}) + \mu_s(\mathbf{r})) M_\rho(\mathbf{\Omega}) M_\nu(\mathbf{\Omega}) d\mathbf{\Omega} \\ &- \big[\sum_{l=0}^{L} \mu_{sl}(\mathbf{r}) \alpha^c_{\rho,l,0} \alpha^c_{\nu,l,0} + 2 \sum_{l=1}^{L} \mu_{sl}(\mathbf{r}) \sum_{m=1}^{l} \alpha^c_{\rho,l,m} \alpha^c_{\nu,l,m} + \alpha^s_{\rho,l,m} \alpha^s_{\nu,l,m} \big],\end{aligned} \tag{3.35}$$

in which, the scattering moments, $\mu_{sl}(\mathbf{r})$ have been truncated to order L. The coefficients $\alpha^c_{\rho,l,m}$ and $\alpha^s_{\rho,l,m}$ are each defined:

$$\alpha^c_{\rho,l,m} = \int_\Omega M_\rho(\mathbf{\Omega}) Y^c_{l,m}(\mathbf{\Omega}) d\mathbf{\Omega}, \qquad \alpha^s_{\rho,l,m} = \int_\Omega M_\rho(\mathbf{\Omega}) Y^s_{l,m}(\mathbf{\Omega}) d\mathbf{\Omega}. \tag{3.36}$$

The scattering/removal operator, $\mathcal{H}$ is therefore described in terms of an absorption coefficient, $\mu_a(\mathbf{r})$, a scattering coefficient, $\mu_s(\mathbf{r})$ and an angularly discretised scattering matrix.

3.4.3 Angular Mass and Jacobian Matrices

The final form of the angularly discretised Boltzmann transport equation is a multi-dimensional differential difference equation representing a symmetric, hyperbolic coupled system:

$$\frac{1}{v}\mathbf{A_t}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial t} + \mathbf{A_x}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial x} + \mathbf{A_y}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial y} + \mathbf{A_z}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial z} + \mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r},t) - \mathbf{S}(\mathbf{r},t) = 0. \tag{3.37}$$

The three $\mathcal{M} \times \mathcal{M}$ angular Jacobian matrices, $\mathbf{A_x}$, $\mathbf{A_y}$, $\mathbf{A_z}$ and the $\mathcal{M} \times \mathcal{M}$ angular mass matrix, $\mathbf{A_t}$ are each defined:

$$\mathbf{A}_t = \int_\Omega M_p(\mathbf{\Omega}) M_\nu(\mathbf{\Omega}) d\mathbf{\Omega} = \mathbf{I}, \qquad \mathbf{A}_k = \int_\Omega M_p(\mathbf{\Omega}) \Omega_k M_\nu(\mathbf{\Omega}) d\mathbf{\Omega}, \tag{3.38}$$

where $\mathbf{A}_k$ is the vector of angular Jacobian matrices $\mathbf{A}_k = (\mathbf{A}_x\ \mathbf{A}_y\ \mathbf{A}_z)^T$ and k = x, y, z. The incremental solid angle $d\mathbf{\Omega} = d\omega d\mu$. The co-ordinate, μ is the co-latitude from the z-axis and ω represents the azimuthal angle. Two angular co-ordinates, defined relative to an orthogonal spatial co-ordinate system, are normally required to specify the direction, $\mathbf{\Omega}$, of multi-dimensional neutral particle travel. If x, y and z are orthogonal spatial co-ordinates, the directional cosines of $\mathbf{\Omega}$ with respect to x, y and z are defined in terms of Ω_x, Ω_y, and Ω_z:

$$\begin{aligned} \Omega_x &= cos(\theta) = \mu \\ \Omega_y &= (1-\mu^2)^{\frac{1}{2}} \cos(\omega) \\ \Omega_z &= (1-\mu^2)^{\frac{1}{2}} \sin(\omega) \end{aligned} \tag{3.39}$$

Only two direction cosines can be specified independently. However -since $\mathbf{\Omega}$ is a unit vector- the third direction cosine can be deduced according to:

$$\Omega_x^2 + \Omega_y^2 + \Omega_z^2 = 1 \tag{3.40}$$

During time-dependent radiation transport calculations, it is desirable to set the angular mass matrix, $\mathbf{A_t}$, equal to the an $\mathcal{M}\times\mathcal{M}$ identity matrix, $\mathbf{I}$. This is possible if orthonormal spherical harmonic or discrete ordinate basis functions are used.

3.4.4 Spatial and Temporal Discretisation Using SUPG Methods

A Streamline Upwind Petrov-Galerkin (SUPG) formulation [58, 59] is used to spatially and temporally discretise the time-dependent Boltzmann transport equation. Prior to discretisation, the angularly discretised Boltzmann transport equation is pre-multiplied by a space-time SUPG term to yield:

$$(\mathbf{I} - \mathbf{A}_{xt} \cdot \nabla_{xt}\mathbf{P}) \left(\frac{1}{v}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial t} + \mathbf{A} \cdot \nabla \mathbf{\Psi}(\mathbf{r},t) + \mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r},t) - \mathbf{S}(\mathbf{r},t) \right) = 0, \quad (3.41)$$

where $\mathbf{P}$ represents a stabilisation matrix, $\mathbf{A} = (\mathbf{A}_x \ \mathbf{A}_y \ \mathbf{A}_z)^T$, $\mathbf{A}_{xt} = (\frac{1}{v}\mathbf{I} \ \mathbf{A})^T$, and $\nabla_{xt} = (\mathbf{I}\frac{\partial}{\partial t} \ \nabla)^T$. The stabilisation matrix is a diagonal matrix of spherical harmonics that determines the amount of weighting assigned to a given finite element discretisation. The stabilisation matrix, is described in more detail in chapter 5. If the second set of brackets on the left hand side of equation 3.41 is defined as the residual. Equation 3.41 may be redefined:

$$(\mathbf{I} - \mathbf{A}_{xt} \cdot \nabla_{xt}\mathbf{P})\mathscr{R}(\mathbf{\Psi}(\mathbf{r},t)) = 0. \quad (3.42)$$

The spatial and temporal dependence of the angularly discretised flux, source and scattering/removal term are discretised in space and time using a finite element approximation:

$$\begin{aligned} \boldsymbol{\psi}(\mathbf{r},t) \approx \mathbf{\Psi}(\mathbf{r},t) &= \sum_{k=1}^{\mathcal{N}} \mathbf{N}_k(\mathbf{r},t)\mathbf{\Psi}_k, \\ \mathbf{s}(\mathbf{r},t) \approx \mathbf{S}(\mathbf{r},t) &= \sum_{k=1}^{\mathcal{N}} \mathbf{N}_k(\mathbf{r},t)\mathbf{S}_k, \\ \boldsymbol{\mathcal{H}}(\mathbf{r}) \approx \mathbf{H}(\mathbf{r}) &= \sum_{j=1}^{\mathcal{N}_H} \mathbf{N}_{H_j}(\mathbf{r})\mathbf{H}_j, \end{aligned} \quad (3.43)$$

in which $\mathbf{N}_k = \mathbf{N}_k(\mathbf{r},t)$ is the amalgamation of all basis functions that are discontinuous in time. $\mathbf{N}_k(\mathbf{r},t)$ and $\mathbf{N}_{H_j}(\mathbf{r})$ are each $(\mathcal{M} \times \mathcal{M})$ diagonal matrices containing finite element basis functions associated with the angular flux, sources and scattering/removal terms respectively. $\mathcal{N}_H$ represents the number of space-time finite element basis functions associated with the materials and $\mathcal{N}$ represents the total number of basis functions over all time steps associated with the angular flux and source terms. $\mathbf{\Psi}_k$, $\mathbf{S}_k$ and $\mathbf{H}_j$ are vectors, of size $\mathcal{M}$, containing the moments of the angularly discretised flux, source

and the scattering/removal terms at nodes k and j respectively. A Petrov-Galerkin finite element method is used to reduce the single Boltzmann transport equation into a series of unknowns. Equation 3.41 is multiplied by a further matrix of finite element basis functions, $\mathbf{N}_i(\mathbf{r},t)$, before the expression is integrated over the space time slab, $V \times \Delta T$. Furthermore, Green's theorem [59] is used to convert the volume integrals representing the angularly discretised streaming terms into surface terms. This ensures that the method is conservative even when reduced or inexact quadrature is used. Incoming boundary conditions, $\mathbf{\Psi}_{in}(\mathbf{r},t)$, on $\mathbf{\Psi}(\mathbf{r},t)$ are thus applied on the surface integral:

$$\begin{aligned}\int_T \int_\Gamma \mathbf{N}_i(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}(\mathbf{r},t) d\Gamma dt &= \int_T \int_{\Gamma_{in}} \mathbf{N}_i(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}_{in}(\mathbf{r},t) d\Gamma dt \\ &+ \int_T \int_{\Gamma_{out}} \mathbf{N}_i(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}(\mathbf{r},t) d\Gamma dt,\end{aligned} \tag{3.44}$$

in which, Γ_{in} represents the incoming flux boundary and Γ_{out} represents the outgoing flux boundary. The boundary conditions associated with the incoming and outgoing flux are approximated using a Riemann approach [58, 61, 62, 129]. The incoming flux is a specified incoming boundary condition and the outgoing flux is taken from the solution internal to the domain, but on the domain boundary, Γ. A Riemann problem is formed normal to the boundary by diagonalising the matrix $\mathbf{A}_s = \mathbf{A}\cdot\mathbf{n}$. That is, using $\mathbf{A}_s = \mathbf{R}_s\mathbf{\Lambda}_s\mathbf{R}_s^{-1}$ and $\mathbf{\Psi}(\mathbf{r},t) = \mathbf{R}_s\mathbf{\Psi}_s(\mathbf{r},t)$ in which, $\mathbf{R}_s$ is a matrix of right hand side eigenvectors, $\mathbf{R}_s^{-1}$ is a matrix of left hand side eigenvectors and $\mathbf{\Lambda}_s$ is a diagonal matrix of eigenvalues associated with the Riemann decomposition. The boundary integral becomes:

$$\int_T \int_\Gamma \mathbf{N}_i(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}(\mathbf{r},t) d\Gamma dt = \int_T \int_\Gamma \mathbf{N}_i(\mathbf{r},t)\mathbf{R}_s\mathbf{\Lambda}_s\mathbf{R}_s^{-1}\mathbf{R}_s\mathbf{\Psi}_s(\mathbf{r},t) d\Gamma dt \tag{3.45}$$

Multiplying equation 3.45 by $\mathbf{R}_s^{-1}$ yields:

$$\int_T \int_\Gamma \mathbf{N}_i(\mathbf{r},t)\mathbf{\Lambda}_s\mathbf{\Psi}_s(\mathbf{r},t)d\Gamma dt = \int_T \int_{\Gamma_{\mathbf{\Lambda}_s>0}} \mathbf{N}_i(\mathbf{r},t)\mathbf{\Lambda}_s\mathbf{\Psi}_s(\mathbf{r},t)d\Gamma dt + \int_T \int_{\Gamma_{\mathbf{\Lambda}_s<0}} \mathbf{N}_i(\mathbf{r},t)\mathbf{\Lambda}_s\mathbf{\Psi}_{s_{in}}(\mathbf{r},t)d\Gamma dt \quad (3.46)$$

When the μ^{th} diagonal component Λ_s^μ of the matrix of eigenvalues $\mathbf{\Lambda}_s$ is positive, this corresponds to an outgoing flux boundary for Riemann variable $\Psi_s^\mu(\mathbf{r},t)$. The solution is then used in the surface integral. When Λ_s^μ is negative, an incoming boundary condition is assumed and the flux condition is placed into the integral. Once the integral is formed it is mapped back into original variables by multiplying equation 3.46 by $\mathbf{R}_s$ and placing it into the fully discretised equation.

Chapter 4

Numerical Inversion

Contents

4.1 Introduction

Two problems are said to be inverses of each other if the formulation of one involves all or part of the solution of the other. The two problems considered, generally referred to as the direct problem and the inverse problem, are obtained from one another by exchanging the roles of the data and the unknowns [77, 78, 79, 147, 148]. The direct (forward) problem represents the information flow specific to a natural, physical process. The direct problem predicts the error free values for the data that correspond to a given model. The inverse problem [77, 78, 79, 147, 148] is the reversal of the sequence, where material properties can be estimated from the known dynamics of the physical system.

During most medical imaging techniques scattering is neglected and the path of the detected photon is assumed to be a straight line between a chosen source-detector pair. For low scattering environments the reconstruction problem can therefore be formulated as a system of linear equations, directly derived from the Boltzmann transport equation. For strongly scattering environments, such as those posed in near infrared optical tomography, the progressive influence of scattering with distance makes the scattering problem markedly more complex and non-linear. The solution therefore does not only involve the calculation of optical properties but must also take in to account scattering effects while inside the medium. Consequently there is not a generally applicable direct method for solving such problems, and instead, inversion techniques are considered to evaluate what material properties would have created the original signal.

Several inversion methods have been developed for imaging problems in highly scattering, heterogeneous media. Such methods include back projection methods, non-linear optimisation methods and perturbation techniques. Non-linear optimisation techniques are of particular interest when considering near infrared optical tomography and are formulated in two parts. Initially the error between the forward model (estimated using trial cross sections) and any measured data is quantified, before the input data is modified, often using a gradient based optimisation method, to reduce the data misfit [71, 72, 73, 149]. Inversion examples using non-linear optimisation techniques are contained in [41, 45, 74, 75, 76, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160].

No general uniqueness results exist for optical tomography because of the range of different unknowns and different measureables. A problem is well-posed if its solution is unique and exists for arbitrary data. Most inversion problems are highly non-linear and ill-posed. That is, the solutions to inversion problems are generally either not unique, not continuous or do not exist for arbitrary data. Generally the existence of a solution to an inverse problem depends on uniqueness, quality of data, and the statistical behaviour of the problem. However, if the problem is non-unique it may be possible to yield a solution if appropriate prior information is available. Prior information may be supplied implicitly in the form of regularisation or explicitly.

Advances in modelling and reconstruction techniques cannot overcome the fact that diffuse optical imaging is a non-unique, ill-posed, underdetermined problem. These detrimental effects may have a profound impact on image quality, particularly spatial resolution. To improve image quality, prior information may be used to reduce the effect of ill-posedness by improving the accuracy of the model. The problem therefore becomes better determined by allowing the small number of measurements obtained to be used in a more effective way. Prior information may be obtained from anatomical imaging techniques or by considering the physics and the physiology of the problem. The most straightforward way to include prior information into the image reconstruction is to use an anatomically realistic forward model. This increases the accuracy of the forward model so that it can represent the measurements more precisely. Further improvements can be made by incorporating anatomical information into the inverse problem and including information about the covariance of the data and the image. A further approach is to express the error functional in a Bayesian framework in which, a statistical expression of the inverse problem may be derived [5, 161].

4.2 Modelling Inversion Using the Adjoint Method

The adjoint model [79, 147, 162, 163] is a tool that has been developed for the inverse modelling of physical systems. The adjoint problem is based on the direct problem and a discrepancy (misfit) calculation between direct model predictions and the measured data.

The discrepancy calculation is contained in an error functional which is minimised using a gradient based optimisation technique [71, 72, 149].

4.2.1 The Formulation of the Direct Model

In operator form, the forward model describes the angular flux, ψ_s, that results from a source placed in region, $\mathbf{r}$ at a boundary location, $\mathbf{r}_s$:

$$\left[\frac{1}{v}\frac{\partial}{\partial t} + \mathbf{\Omega}\cdot\boldsymbol{\nabla} + \mathcal{H}\right]\psi_s = C_s\delta(\mathbf{r}-\mathbf{r}_s)\delta(t). \tag{4.1}$$

Equation 4.1 describes a series of $\mathcal{S}$ source problems, where $s = (1, 2, ..., \mathcal{S})$. The source is a Dirac-delta function in time, $\delta(t)$, and space, $\delta(\mathbf{r}-\mathbf{r}_s)$ and is centred at position $\mathbf{r}_s$ of the domain $\mathbf{r}$ at time t. The magnitude of the source strength is given by C_s. ψ_s represents the angular flux owing to the source number, s, while the scattering/removal operator, $\mathcal{H}$, contains the spatially varying material properties. Each source problem yields an angular flux solution at all points in the solution domain. If the source position is located at a node then the source strength and angular flux are equal to the finite element basis function centred on that node multiplied by a scalar (to scale the strength).

$$\mathbf{b}_{s_i} = \int_T\int_V \mathbf{N}_i(\mathbf{r},t)\mathbf{S}_s dV dt, \tag{4.2}$$

in which V represents the solution domain and T represents the time domain. $\mathbf{b_s} = (\mathbf{b_s}_{\,1\,1}\ \mathbf{b_s}_{\,1\,2}\ ...\ \mathbf{b_s}_{\mathcal{N}\mathcal{M}})^T$, contains the discretised source contributions and $\mathbf{S}_s$ is a finite element approximation of the source. If the source is located inside an element it must be distributed to the nearest node associated with that element and possibly nodes associated with neighbouring elements. Following discretisation, the direct problem is defined according to the matrix equation:

$$\mathbf{E}\mathbf{\Psi}_s = \mathbf{b_s}. \tag{4.3}$$

Each source problem results in a solution vector for the angular flux. For $s = (1, 2, ..., \mathcal{S})$ the angular flux is defined $\mathbf{\Psi}_s = (\mathbf{\Psi_s}_{1\,1}\ \mathbf{\Psi_s}_{1\,2}\ ...\ \mathbf{\Psi_s}_{\mathcal{N}\mathcal{M}})^T$. Matrix $\mathbf{E}$ depends on the discretisation method that is used and has a block sparse structure. By solving the discrete linear system described in equation 4.3, a solution for the angular flux at all points along the finite element mesh is found. The GMRES (Genralised Residual Method) iterative solver is used to solve the discrete linear system [164] described in equation 4.3. The GMRES algorithm is a Kyrlov type solver which can generate an approximation to any general system by iteratively searching along the directions defined by the Krylov space:

$$K_m(\mathbf{E}, \mathbf{r}_0) = \{\mathbf{r}_0, \mathbf{E}\mathbf{r}_0, \mathbf{E}^2\mathbf{r}_0, \ldots, \mathbf{E}^{m-1}\mathbf{r}_0\}. \tag{4.4}$$

Vector $\mathbf{r}_0$ represents the initial residual of an initial estimate given by $\mathbf{\Psi}_{s_0}$ The scalar m defines the size of the search space (i.e. the number of search directions). During each iterative update the approximation $\mathbf{\Psi}_s$ represents the projection onto the Krylov sub-space that finds the vector with the minimum residual. That is, GMRES finds $\mathbf{\Psi}_s$ that satisfies,

$$\min_{\mathbf{\Psi}_{s_i}^{'} \in \mathbf{\Psi}_{s0}+K_i} \|\mathbf{b}_s - \mathbf{E}\mathbf{\Psi}_{s_i}^{'}\|, \tag{4.5}$$

where $\|.\|$ defines the 2-norm. GMRES incorporates block forward and backward Gauss Siedel preconditioning. Each block is an $\mathcal{M} \times \mathcal{M}$ submatrix associated with each node of the finite element mesh.

4.2.2 The Formulation of the First Order Adjoint Method

The adjoint method generally exists in the same form as the direct problem but from the viewpoint of the system's control variables [79]. The control variables often take the form of boundary conditions or the material properties associated with the medium. Providing the initial model is close enough to the global solution to allow successive relinearisations, the model trajectory can be adjusted as close as possible to the measured data. Using the product rule [91, 92], the variation due to the control variables may be expressed as:

$$\frac{\partial \mathbf{E}}{\partial m_j}\,\mathbf{\Psi}_s + \mathbf{E}\frac{\partial \mathbf{\Psi}_s}{\partial m_j} = \frac{\partial \mathbf{b}_s}{\partial m_j}. \tag{4.6}$$

where, $\mathbf{m} = (m_1\ m_2\ ...\ m_{\mathcal{N}_H})^T$, represents the system's control variables. In medical optical tomography $\mathbf{m}^T = (\mathbf{m}_a^T\ \mathbf{m}_{s0}^T)$, in which $\mathbf{m}_a = (m_{a1}\ m_{a2}\ ...\ m_{a\mathcal{N}_H})^T$ represents the discretised absorption material properties where -assuming m_{a_i} is the nodal value of a- $m_{a_i} = ln(\mu_{a_i})$ and $\mathbf{m}_{s0} = (m_{s01}\ m_{s02}\ ...\ m_{s0\mathcal{N}_H})^T$ represents the discretised scattering material properties in which $m_{s0_i} = ln(\mu_{s0_i})$. $\mathbf{m}_a$ and $\mathbf{m}_{s0}$ are approximated using a set of finite element basis functions, N_{H_j}, such that:

$$m_a = \sum_{j=1}^{\mathcal{N}_H} N_{H_j} m_{a_j}, \qquad m_{s0} = \sum_{j=1}^{\mathcal{N}_H} N_{H_j} m_{s0_j}. \tag{4.7}$$

in which, $j = (1, 2, ..., \mathcal{N}_H)$. Material properties are considered as opposed to material cross-sections because the resulting contributions ensure positivity. The error functional is optimised using standard least squares techniques and may result in uniform material properties. In addition, material cross-sections, sometimes vary by several orders of magnitude, which is undesirable when considering model regularisation. The error functional becomes a negative log of a probability distribution assuming Gaussian noise.

$$F_d = \frac{1}{2}\sum_{s=1}^{\mathcal{S}}\sum_{r=1}^{\mathcal{R}}\sum_{\tau=1}^{\mathcal{T}}\sum_{\mu=1}^{\mathcal{M}}\sum_{i=1}^{\mathcal{N}} W^r_{s\ i}\left(\Psi_{s\ \tau\ \mu\ i} - d^r_{s\ \tau\ \mu}\right)^2. \tag{4.8}$$

Equation 4.8 represents the weighted sum of the squared differences between observed data, $d^r_{s\ \tau\ \mu}$, and data predicted from a given material property distribution, $\Psi_{s\ \tau\ \mu\ i}$, for all $\mathcal{S}$ sources and $\mathcal{R}$ detectors. The summation is also performed over each node, $i = (1, 2, ..., \mathcal{N})$, for each time step, $\tau = (1, 2, ..., \mathcal{T})$ and for each degree of freedom, $\mu = (1, 2, ..., \mathcal{M})$. The weighting factor, $W^r_{s\ i}$, distributes both the sources and detectors to the nearest nodes using a Gaussian approximation (see chapters 5 and 6). The error functional is minimised using an iterative, gradient based, optimisation algorithm [71, 72]. Starting with an initial guess, for each iteration an improved vector of variables is searched. The search direction is computed from the gradient of the error functional with respect to the

control variables. Using the chain rule [91, 92] :

$$\frac{\partial F}{\partial m_j} = \left(\frac{\partial \mathbf{\Psi}_s}{\partial m_j}\right)^T \frac{\partial F}{\partial \mathbf{\Psi}_s} \tag{4.9}$$

Re-arranging equation 4.6

$$\frac{\partial \mathbf{\Psi}_s}{\partial m_j} = -\mathbf{E}^{-1}\left(\frac{\partial \mathbf{E}}{\partial m_j}\mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial m_j}\right) \tag{4.10}$$

and transposing equation 4.10

$$\frac{\partial \mathbf{\Psi}_s}{\partial m_j}^T = -\left(\frac{\partial \mathbf{E}}{\partial m_j}\mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial m_j}\right)^T \mathbf{E}^{-T} \tag{4.11}$$

Substituting equation 4.11 in to equation 4.9 the resulting gradient of the error functional can be found

$$\frac{\partial F}{\partial m_j} = -\left(\frac{\partial \mathbf{E}}{\partial m_j}\mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial m_j}\right)^T \cdot \left(\mathbf{E}^{-T}\frac{\partial F}{\partial \mathbf{\Psi}_s}\right) \tag{4.12}$$

Defining $\mathbf{E}^{-T}\dfrac{\partial F}{\partial \mathbf{\Psi}_s}$ = $\mathbf{\Psi}_s^*$, the adjoint equation can be defined

$$\mathbf{E}^T\mathbf{\Psi}_s^* = \frac{\partial F}{\partial \mathbf{\Psi}_s} \tag{4.13}$$

$\mathbf{\Psi}_s^* = (\mathbf{\Psi}^*_{\mathbf{s}\,1\,1}\ \mathbf{\Psi}^*_{\mathbf{s}\,1\,2}\ ...\ \mathbf{\Psi}^*_{\mathbf{s}\,\mathcal{N}\mathcal{M}})^T$, represents the vector of (adjoint) control variables.

4.3 Gradient Based Optimisation Methods

Least square methods assume a Gaussian approximation for the distribution of a random variable [71, 149]. To obtain the best fit between a model estimation and observational data, the minimum of the error functional is sought. This occurs when the gradient of the

error functional, F, is equal to zero. The best fit measurements can then be incorporated into an iterative search to find a solution to a given problem [71, 72, 165, 166]. Popular optimisation techniques that are discussed include the Steepest Descent method [71, 72], the Non-Linear Conjugate Gradient method [71, 72] and the second order Levenberg-Marquardt method [71, 72, 73].

4.3.1 The Steepest Descent Method

From an initial starting point, $\mathbf{m}_0$, the Steepest Descent (also known as Gradient Descent) [71, 72] method updates that estimate by taking a step in the direction where the gradient of the error functional decreases most rapidly. This is commonly referred to as the negative gradient direction:

$$\mathbf{a} = -\frac{\partial F}{\partial \mathbf{m}}, \tag{4.14}$$

where $\mathbf{a} = (a_1\ a_2\ ...\ a_{\mathcal{NM}})^T$. The control variables are then updated according to:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k \mathbf{a}, \tag{4.15}$$

where α_k is a carefully chosen constant. Providing α_k is chosen correctly, a monotonic decrease in the error functional is ensured and a minimum eventually reached. However -since α_k is usually a relatively small number- the rate of convergence is often very slow.

4.3.2 The Gauss-Newton Method

The Gauss-Newton method [71, 72, 167, 158] is based on the assumption that the error functional can be accurately represented by a quadratic approximation at point $\mathbf{m}_0$. A Taylor Series expansion [91, 92] of the error functional can therefore be truncated to include only first and second order derivatives:

$$F(\mathbf{m}_0 + \Delta\mathbf{m}) = F(\mathbf{m}_0) + \mathbf{a}^T\Delta\mathbf{m} + \frac{1}{2}(\Delta\mathbf{m}^T\mathbf{B}\Delta\mathbf{m}). \tag{4.16}$$

$\mathbf{a}$, represents a gradient vector of size $\mathcal{NM}$ and $\mathbf{B}$ is an $(\mathcal{NM} \times \mathcal{NM})$ Hessian matrix of second order derivatives. The Hessian matrix measures the curvature of the error surface. The minimum of the error functional occurs when the gradient of the error functional is equal to zero (which is often achieved using linear conjugate gradient methods). At this point the truncated Taylor Series is reduced to:

$$\mathbf{a} + \mathbf{B}\Delta\mathbf{m} = 0, \tag{4.17}$$

Re-arranging equation 4.17 yields an expression for the incremental parameter change:

$$\Delta\mathbf{m} = -\mathbf{B}^{-1}\mathbf{a}, \tag{4.18}$$

where the gradient vector and Hessian matrix are respectively defined as:

$$\mathbf{a}_j = \frac{\partial F}{\partial m_j} = \sum_{i=1}^{\mathcal{N}} \frac{1}{\sigma^2}\epsilon_i\frac{\partial \epsilon_i}{\partial m_j}, \tag{4.19}$$

$$\mathbf{B}_{j,k} = \frac{\partial^2 F}{\partial m_j \partial m_k} = \sum_{i=1}^{\mathcal{N}} \frac{1}{\sigma^2}\frac{\partial \epsilon_i}{\partial m_j}\frac{\partial \epsilon_i}{\partial m_k} + \epsilon_i\frac{\partial^2 \epsilon_i}{\partial m_j \partial m_k}, \tag{4.20}$$

and vector, $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_0\ \boldsymbol{\epsilon}_1\ ....\ \boldsymbol{\epsilon}_{\mathcal{NM}})$, is a residual vector. For sufficiently small residuals, the second term in equation 4.20 is ignored, resulting in an Approximate Hessian matrix. Defining the Jacobian matrix, $\mathbf{J}$, as the gradient of the residual with respect to the control variables allows the gradient vector and Hessian matrix to de expressed in their vector form:

$$\mathbf{a} = -\mathbf{J}^T\mathbf{W}\boldsymbol{\epsilon}, \tag{4.21}$$

$$\mathbf{B} = \mathbf{J}^T\mathbf{W}\mathbf{J}. \tag{4.22}$$

The parameter change, $\Delta\mathbf{m}$, and parameter update, $\mathbf{m}_{k+1}$, are respectively defined:

$$\Delta\mathbf{m} = \mathbf{B}^{-1}\mathbf{a} = (\mathbf{J}^T\mathbf{W}\mathbf{J})^{-1}(-\mathbf{J}^T\mathbf{W}\boldsymbol{\epsilon}), \tag{4.23}$$

$$\mathbf{m}_{k+1} = \mathbf{m}_k + (\mathbf{J}^T\mathbf{W}\mathbf{J})^{-1}(-\mathbf{J}^T\mathbf{W}\boldsymbol{\epsilon}). \tag{4.24}$$

Providing an initial iterate is sufficiently close to the solution, the Gauss-Newton method has a quadratic rate of convergence. Divergence may occur if a poor starting vector (i.e. a vector far from the minimum) is chosen because of the neglected second term in the Hessian matrix. The Gauss-Newton method always fails if the Hessian becomes singular or ill-conditioned.

4.3.3 The Levenberg-Marquardt Method

The Gauss-Newton method can be modified so that convergence always occurs. Far from the minimum the Gauss-Newton method is not reliable. Near the minimum the Gauss-Newton method converges very quickly while the Steepest Descent method converges more slowly. The Levenberg-Marquardt method [71, 72, 73, 168] uses the Steepest Descent method far from a solution's minimum while switching to the Gauss-Newton method once in the vicinity of the minimum. The approximate Hessian is replaced with [74]:

$$\mathbf{B} = \mathbf{J}^T\mathbf{W}\mathbf{J} + \lambda\mathbf{M}_{H_L}, \tag{4.25}$$

where $\mathbf{M}_{H_L}$ is the diagonal mass matrix and is found using finite element basis functions, N_{H_i}, according to, $\mathbf{M}_{H_{L_{ii}}} = \int_V N_{H_i} dV$. λ is a conditioning factor that controls the interpolation between the Steepest Descent and Gauss-Newton methods. $\lambda\mathbf{M}_{H_L}$ controls

the magnitude of the model update $\Delta\mathbf{m}$. If λ is large then $\Delta\mathbf{m}$ stays small. In addition, $\lambda\mathbf{M}_{H_L}$ aids the conditioning of the Hessian matrix by increasing its diagonal dominance. The iteration update is defined according to:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + (\mathbf{J}^T\mathbf{W}\mathbf{J} + \lambda_k\mathbf{M}_{H_L})^{-1}(-\mathbf{J}^T\mathbf{W}\boldsymbol{\epsilon}). \tag{4.26}$$

A suitably large starting value of λ must be given. As λ tends to 0 the search direction approaches the Gauss-Newton method, and as λ tends to infinity the search direction reduces to the Steepest Descent method. λ is usually decreased by a factor of 10 when iterates are successful and increased by a factor of 10 when the iteration fails.

4.3.4 Conjugate Gradient Methods

Although the error functional decreases most rapidly along the negative gradient direction, it does not necessarily produce the fastest convergence. Conjugate gradient algorithms [71, 72] are gradient methods that perform searches along conjugate directions. Conjugate directions generally produce faster convergence than Steepest Descent directions. A variable step size -which is used to minimise the error functional- is computed and adjusted at each iteration point. All Conjugate Gradient algorithms begin by searching in the Steepest Descent direction.

The Preconditioned Linear Conjugate Gradient Method

The Linear Conjugate Gradient method [72, 74] is an iterative scheme that is used to solve linear systems with positive definite coefficient matrices. The distribution of eigenvalues in a given coefficient matrix contributes to the performance of an algorithm and can be made more favourable by preconditioning the linear system. The preconditioner [72], $\hat{\mathbf{B}}$, is an approximation of the Hessian matrix, $\mathbf{B}$, but is easier to invert. The preconditioning matrix can, be calculated in a number of different ways [72] and must approximate $\mathbf{B}$ well enough to improve convergence after each iteration. Like the Hessian matrix, matrix, $\hat{\mathbf{B}}$ is

non-singular and symmetric positive definite. For the convex quadratic Taylor's expansion described in equation 4.16, the gradient of the error functional results in some residual:

$$\mathbf{r}_0 = \mathbf{a} - \mathbf{B}\Delta\mathbf{X}_0. \tag{4.27}$$

An initial starting point is defined in terms of the residual, $\mathbf{r}$, and the preconditioner:

$$\mathbf{m}_0 = \hat{\mathbf{B}}^{-1}\mathbf{r}_0. \tag{4.28}$$

A set of non-zero vectors $\mathbf{r} = (\mathbf{r}_0\ \mathbf{r}_1\ ....\ \mathbf{r}_{\mathcal{NM}})$ are conjugate with respect to a symmetric, positive definite preconditioner, $\hat{\mathbf{B}}^{-1}$, if:

$$\mathbf{r}_k^T\hat{\mathbf{B}}^{-1}\mathbf{r}_k = 0. \tag{4.29}$$

The error functional is minimised in L steps (or less) along the individual directions of the conjugate set. Starting with an initial point, $\mathbf{X}_0$, and a set of conjugate directions, a sequence is generated by setting:

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \alpha_k\mathbf{m}_K, \tag{4.30}$$

where α_k is defined in terms of the conjugacy condition:

$$\alpha_k = \frac{\mathbf{r}_k^T\hat{\mathbf{B}}^{-1}\mathbf{r}_k}{\mathbf{m}_k^T\mathbf{B}\mathbf{m}_k}. \tag{4.31}$$

If α_k is close to 1 then the preconditioner is considered a close approximation of the Approximate Hessian given in equation 4.22. The solver is assumed to be converging well if α_k is greater than 0.01 on the final linear iteration. The residual, $\mathbf{r}$, is updated according to:

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{B}\mathbf{m}_k. \tag{4.32}$$

The new parameter update, $\mathbf{m}_{k+1}$, is finally calculated from a combination of the previous vector, $\mathbf{m}_k$ and the current descent direction.

$$\mathbf{m}_{k+1} = \hat{\mathbf{B}}^{-1}\mathbf{r}_{k+1} + \beta_k \mathbf{m}_k, \tag{4.33}$$

where β_k requires that $\mathbf{r}_k$ and $\mathbf{r}_{k+1}$ are conjugate with respect to $\mathbf{B}$. An expression for β_k is therefore found by premultiplying equation 4.33 by $\mathbf{r}_{K+1}^T$ and imposing the conjugacy condition.

$$\beta_k = \frac{\mathbf{r}_{k+1}^T \hat{\mathbf{B}}^{-1}\mathbf{r}_{k+1}}{\mathbf{r}_k^T \hat{\mathbf{B}}^{-1}\mathbf{r}_k}. \tag{4.34}$$

Thus, for linear calculations, a preconditioned linear conjugate gradient solver is used to perform successive one dimensional minimisations along each of the conjugate search directions.

The Non-Linear Conjugate Gradient Method

The Linear Conjugate Gradient method [71, 72] can be adapted to minimise general convex functions or general non-linear functions. Instead of computing a value for α_k (given in equation 4.31) a line search is performed, that identifies an approximate minimum of the non-linear error functional along $\mathbf{m}_k$. The residual is then replaced with the gradient of the non-linear error functional. The search direction for each iteration is defined:

$$\mathbf{m}_{k+1} = -\mathbf{a}_{k+1} + \beta_k \mathbf{m}_k, \tag{4.35}$$

in which, β_k can be calculated according to different techniques. Two of the most popular methods used to calculate β_k are the Polak-Ribere update and the Fletcher Reeves update.

The Polak-Ribiere update - β_k is calculated according to the inner change of the previous change in the gradient with the current gradient divided by the norm squared of the previous gradient:

$$\beta_k = \frac{\Delta \mathbf{a}_{k-1}^T \mathbf{a}_k}{\mathbf{a}_{k-1}^T \mathbf{a}_{k-1}} \tag{4.36}$$

The Fletcher-Reeves update - β_k is the ratio of the norm squared of the current gradient to the norm squared of the previous gradient.

$$\beta_k = \frac{\Delta \mathbf{a}_k^T \mathbf{a}_k}{\mathbf{a}_{k-1}^T \mathbf{a}_{k-1}} \tag{4.37}$$

The Polak-Riebiere method tends to be the more efficient and robust algorithm and has therefore been incorporated into the Non-Linear Conjugate Gradient algorithm used in our code. Non-linear Conjugate Gradient methods reduce to Linear Conjugate Gradient methods when the error functional is a strongly convex quadratic function and the line search is exact. The search direction is periodically reset to the negative gradient when the number of iterations equals the number of network parameters.

The Line Search - Brents Search

The non-linear conjugate gradient method generally requires a line search to be performed to calculate β_k. The Brents Search is a hybrid combination of a Golden Section Search [72] and a quadratic interpolation. The Golden Section Search evaluates the optimisation performance at a given distance and increases that distance by a specified amount at each step along the search direction. When the performance increases between two iterations a minimum is bracketed and a series of points are located within the initial interval. A quadratic function is then fitted to these points and the minimum of the quadratic function is computed. Providing the minimum falls within an appropriate level of uncertainty, the performance of the individual points determines the section of the interval that can be discarded. A new interior point is then placed within the new interval and a new minimum is found. If the minimum falls outside the uncertainty then a new Golden Section Search is performed.

Chapter 5

A Finite Element Steady-State Optical Imaging Method for Optically Thick and Optically Thin Media

Contents

The nature of optical tomography inversion is to reconstruct cross-sectional images of the optical properties of highly scattering biological tissue from source measurements and detector readings performed on the surface of the domain. In this chapter, the transport of optical radiation is described by the one-speed, steady state Boltzmann transport equation. Using a full transport theory forward model to predict light propagation enables accurate modelling of transparent, void-like regions, which are commonly found in biological media. A Stream-line Upwinded Petrov-Galerkin (SUPG) finite element method is used to discretise the Boltzmann transport equation and spherical harmonic basis functions are used to represent the angle of photon travel. A stabilisation term designed specifically for inverse problems is incorporated into the discretisation to ensure an optimal modelling method for both optically thick and optically thin media.

An error functional representing the forward model misfit measured at the detectors is minimised using a gradient based, least squares optimisation scheme and is obtained by differentiating the discrete forward model. This results in the reconstruction of the spatial distribution of scattering and absorption coefficients inside the domain. Photon propagation in layered tissue environments can produce non-unique solutions to the inverse problem when insufficient data is supplied to the model. Special attention is therefore given to the use of model covariance matrices and data weighting to assist the inversion process to arrive at a physically plausible result.

5.1 Introduction

Transport theory provides a deterministic numerical method that describes the migration of neutral particles according to the Boltzmann transport equation. In this chapter a steady state form of the mono-energetic Boltzmann transport equation have been used to model the distribution of optical radiation inside a given medium. Prior to inversion a direct numerical discretisation scheme, subdivides the solution domain into a finite number of angular, ($\mathbf{\Omega}$ = (θ, ω), and spatial ($\mathbf{r} = (x, y, z)$), components. Spatial discretisation is achieved using a Streamline Upwinded Petrov-Galerkin (SUPG) formulation and the angle of photon travel is described according to spherical harmonics. The SUPG method has been designed specifically for inversion regimes and yields a set of discrete equations that can be differentiated with respect to an object's material properties. This allows gradients to be formed as part of the inversion procedure. The SUPG method also incorporates an exponential stabilisation matrix, in order to provide the numerical dissipation necessary to model biological media. The exponential stabilisation matrix provides an adequate amount of weighting to model both optically thick (diffusive) and transparent media, media with highly scattering and absorbing material properties and media with abruptly changing material properties.

Scattering dominant light transport [63, 64, 65] is a limiting factor in the development of optical tomography. During most conventional medical imaging techniques (such as CT and SPECT) scattering is neglected and the path of the detected photon is assumed to be a straight line between a chosen source-detector pair. For strongly scattering environments, such as those posed in near infrared optical tomography, the progressive influence of scattering with distance makes the scattering problem complex and non-linear. There is not a generally applicable direct method for solving highly scattering media, and instead, inversion methods are applied to evaluate optical properties. A multi-dimensional, non-linear inversion scheme is developed here to describe the transport of photons during near infrared optical tomography. The technique enables the reconstruction of optical property distribution in two and three dimensions for arbitrary model geometries and for arbitrary source and detector positions along the boundary of the domain. The inverse problem is posed as a functional optimisation problem [69, 41, 74, 75, 76]. Functional optimisation

techniques employ a forward model that provides detector readings based on estimates of the distribution of optical properties inside the medium. The predicted detector readings are compared alongside observed data and are quantified using an error functional. The functional is minimised by iteratively updating the trial estimates and performing new forward calculations with these updated optical properties. Iterations are performed until the predicted data agrees with the detector readings. The final distribution of optical properties is then displayed as an image.

Most inversion problems are highly non-linear and ill-posed [74, 75]. That is, the solutions to inversion problems are generally either not unique, not continuous or do not exist for arbitrary data. Model covariance regularisation overcomes the problem of ill-posedness at the expense of limiting the allowed models to a class of model that is compatible with the provided model covariance information. Functional optimisation techniques are highly desirable because they allow the straightforward inclusion of model covariance constraints by including additional terms in the error functional. Two additional functional terms have been added to the data misfit functional, penalising both structure and any deviation from the starting model [74, 75]. It is also desirable to use second order derivative optimisation schemes such as the Levenberg-Marquardt method over first order optimisation schemes such as the non-linear conjugate gradient method, since Levenberg-Marquardt calculations treat regularisation terms implicitly. Although Levenberg-Marquardt methods involve computationally time consuming calculations involving both Jacobian and approximate Hessian matrices, the assembly and storage of the Jacobian and approximate Hessian matrix may be bypassed alongside the need to implicitly invert the approximate Hessian. This is achieved using preconditioned conjugate gradient (PCG) solvers to iteratively solve the associated least squares problem.

The following chapter is organised as follows. Section 5.2 describes the basic equations used to model steady state radiation transport alongside their finite element discretisation. Section 5.3 defines the error functional used in the inversion of non-linear ill-posed problems. The error functional has terms representing data misfit and model covariance with a view of penalising structure and any deviation from the starting model. The optimisation of the error functional by means of a modified Levenberg-Marquardt method and a description of how calculations involving Jacobian and approximate Hessian matrices are

approached is described in section 5.4 and 5.5, while the numerical applications described in section 5.6 demonstrate the feasibility of the proposed method. Finally, our findings are summarised in section 5.7.

5.2 Governing Equations

For optical tomography applications the steady state Boltzmann transport equation is used to describe the angular flux, $\psi(\mathbf{r}, \mathbf{\Omega})$, that results from an external source of photons, $s(\mathbf{r}, \mathbf{\Omega})$, incident on a host medium. The steady state Boltzmann transport equation for mono-energetic photons is written:

$$\mathbf{\Omega} \cdot \nabla\psi(\mathbf{r}, \mathbf{\Omega}) + \mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}) = s(\mathbf{r}, \mathbf{\Omega}), \tag{5.1}$$

where $\mathbf{r}$ is the position vector $\mathbf{r} = (x, y, z)$ of a particle or photon moving along direction, $\mathbf{\Omega} = (\theta, \omega)$. $\psi(\mathbf{r}, \mathbf{\Omega})$ is the particle angular flux and $s(\mathbf{r}, \mathbf{\Omega})$ represents an external source term. The scattering/removal operator, $\mathcal{H}$ is defined:

$$\mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}) = (\mu_a(\mathbf{r}) + \mu_s(\mathbf{r}))\psi(\mathbf{r}, \mathbf{\Omega}) - \int_{\Omega} \mu_s(\mathbf{r}, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega})\psi(\mathbf{r}, \mathbf{\Omega}^{'}) d\mathbf{\Omega}^{'},$$

where $\mu_a(\mathbf{r})$ and $\mu_s(\mathbf{r})$ represent position dependent absorption and scattering coefficients respectively, and $\mu_s(\mathbf{r}, \mathbf{\Omega}^{'} \rightarrow \mathbf{\Omega})$ represents the differential macroscopic cross-section. To complete the mathematical description of the Boltzmann transport equation the boundary conditions must be defined for a particular problem. Each of the models presented in this chapter have vacuum boundary conditions [34]:

$$\psi(\mathbf{r}_s, \mathbf{\Omega}) = 0, \quad when \quad \mathbf{n} \cdot \mathbf{\Omega} < 0. \tag{5.2}$$

$\mathbf{n}$ represents the outward facing normal to the boundary.

5.2.1 The Finite Element Approximation of the Forward Model

In order to solve equation 5.1 numerically the source term is subtracted from both sides of equation 5.1 and the subsequent expression is discretised in angle and space. Angular discretisation is achieved using the Petrov-Galerkin method [58, 59, 119, 120]. The Petrov-Galerkin method is applied by multiplying the entire expression by a series of angularly dependent finite element weight functions, $\mathcal{M}_\rho(\mathbf{\Omega})$, and integrating the subsequent expression over the solution domain. Additionally a finite element approximation $\mathbf{\Psi}(\mathbf{r})$, replaces the continuous angular flux, $\psi(\mathbf{r}, \mathbf{\Omega})$ and a finite element approximation, $\mathbf{S}(\mathbf{r})$, is also used to approximate the source function, $\mathbf{s}(\mathbf{r}, \mathbf{\Omega})$. The system of equations, for steady state radiation transport becomes:

$$\int_{\mathbf{\Omega}} M_\rho(\mathbf{\Omega}) \left[\mathbf{\Omega} \cdot \nabla\psi(\mathbf{r}, \mathbf{\Omega}) + \mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}) - s(\mathbf{r}, \mathbf{\Omega})\right] d\mathbf{\Omega} = 0. \tag{5.3}$$

The approximation $\mathbf{\Psi}(\mathbf{r})$ for the angular flux, $\psi(\mathbf{r}, \mathbf{\Omega})$, is described using the finite element basis function, $M_\nu(\mathbf{\Omega})$:

$$\psi(\mathbf{r}, \mathbf{\Omega}) \approx \sum_{\nu=1}^{\mathcal{M}} M_\nu(\mathbf{\Omega})\Psi_\nu(\mathbf{r}), \tag{5.4}$$

where $\mathbf{\Psi}(\mathbf{r}) = (\Psi_1(\mathbf{r})\ \Psi_2(\mathbf{r})\ ...\ \Psi_\mathcal{M}(\mathbf{r}))^T$ represents a vector of $\mathcal{M}$ angular moments and $M_\nu(\mathbf{\Omega})$ represents the $\mathcal{M}$ (spherical harmonic) angular basis functions. Likewise, the approximation, $\mathbf{S}(\mathbf{r})$, for the source function $s(\mathbf{r}, \mathbf{\Omega})$, is also expanded using a discrete series of finite element basis functions:

$$s(\mathbf{r}, \mathbf{\Omega}) \approx \sum_{\nu=1}^{\mathcal{M}} M_\nu(\mathbf{\Omega})S_\nu(\mathbf{r}), \tag{5.5}$$

where $\mathbf{S}(\mathbf{r}) = (S_1(\mathbf{r})\ S_2(\mathbf{r})\ ...\ S_\mathcal{M}(\mathbf{r}))^T$. For practical applications it is often assumed that scattering collisions have rotational symmetry. Therefore the differential scattering coefficient, $\mu_s(\mathbf{r},\mathbf{\Omega}' \rightarrow \mathbf{\Omega})$, is not dependent on the incident and scattered particle directions, but instead on the change in particle direction, μ_0. Consequently, $\mu_s(\mathbf{r}, \mu_0)$ may be

expanded as a series of orthogonal Legendre Polynomials [112]. Following expansion, Legendre addition theorem [112] is used to expand each orthogonal Legendre polynomial as a series of directionally dependent spherical harmonics [32, 34, 59, 112, 109], thus simplifying the scattering/removal matrix $\mathcal{H}$. The angular discretisation of the scattering/removal operator, $\mathcal{H}$, yields the diagonal matrix (for spherical harmonic discretisations):

$$\begin{aligned} \boldsymbol{H}_{\rho\nu}(\mathbf{r}) &= \int_{\mathbf{\Omega}} (\mu_a(\mathbf{r}) + \mu_s(\mathbf{r})) M_\rho(\mathbf{\Omega}) M_\nu(\mathbf{\Omega}) d\mathbf{\Omega} \\ &- \Big[\sum_{l=0}^{L} \mu_{sl}(\mathbf{r}) \alpha^c_{\rho,l,0} \alpha^c_{n,l,0} + 2 \sum_{l=1}^{L} \mu_{sl}(\mathbf{r}) \sum_{m=1}^{l} \alpha^c_{\rho,l,m} \alpha^c_{\nu,l,m} + \alpha^s_{\rho,l,m} \alpha^s_{\nu,l,m} \Big] \\ &. \end{aligned} \tag{5.6}$$

The scattering moments, $\mu_{sl}(\mathbf{r})$ are derived from the orthogonality of the Legendre polynomials and the expansion has been truncated to order L. The coefficients $\alpha^c_{\rho,l,m}$ and $\alpha^s_{\rho,l,m}$ are each given by:

$$\alpha^c_{\rho,l,m} = \int_{\Omega} M_\rho(\mathbf{\Omega}) Y^c_{l,m}(\mathbf{\Omega}) d\mathbf{\Omega}, \quad \alpha^s_{\rho,l,m} = \int_{\Omega} M_\rho(\mathbf{\Omega}) Y^s_{l,m}(\mathbf{\Omega}) d\mathbf{\Omega}. \tag{5.7}$$

Furthermore, $Y^c_{l,m}(\mathbf{\Omega})$, and $Y^s_{l,m}(\mathbf{\Omega})$ represent the real and imaginary parts of the spherical harmonic function, $Y_{l,m}(\mathbf{\Omega})$. If $P_{l,m}(\mu)$ denotes associated Legendre polynomials:

$$Y_{l,m}(\mathbf{\Omega}) = \begin{cases} Y^c_{l,m}(\mathbf{\Omega}) = \sqrt{\dfrac{(2l+1)(l-m)!}{(l+m)!}}\, P_{l,m}(\mu)\, cos(m\omega), \;\; 0 \le m \le l \\ Y^s_{l,m}(\mathbf{\Omega}) = \sqrt{\dfrac{(2l+1)(l-m)!}{(l+m)!}}\, P_{l,m}(\mu)\, sin(m\omega), \;\; 0 \le m \le l. \end{cases}$$

The final form of the angularly discretised Boltzmann transport equation is a multi-dimensional differential difference equation representing a symmetric, hyperbolic coupled system:

$$\mathbf{A_x}\frac{\partial \mathbf{\Psi}(\mathbf{r})}{\partial x} + \mathbf{A_y}\frac{\partial \mathbf{\Psi}(\mathbf{r})}{\partial y} + \mathbf{A_z}\frac{\partial \mathbf{\Psi}(\mathbf{r})}{\partial z} + \mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r}) - \mathbf{S}(\mathbf{r}) = 0, \tag{5.8}$$

The three $\mathcal{M} \times \mathcal{M}$ angular Jacobian matrices, $\mathbf{A_x}, \mathbf{A_y}, \mathbf{A_z}$ are given by:

$$\mathbf{A}_{k_{\rho\nu}} = \int_{\Omega} M_{\rho}(\mathbf{\Omega})\Omega_k M_{\nu}(\mathbf{\Omega}) d\mathbf{\Omega}, \tag{5.9}$$

where $\mathbf{A}_k$ is the vector of angular Jacobian matrices $\mathbf{A}_k = (\mathbf{A}_x\ \mathbf{A}_y\ \mathbf{A}_z)^T$ and $k = x, y, z$. The incremental solid angle $d\mathbf{\Omega} = d\omega d\mu$. The co-ordinate μ is the co-latitude from the z-axis and ω represents the azimuthal angle. Two angular co-ordinates, defined relative to an orthogonal spatial co-ordinate system, are normally required to specify the direction, Ω_k, of multi-dimensional neutral particle travel. If x, y and z are orthogonal spatial co-ordinates, the directional cosines of Ω_k with respect to x, y and z are defined in terms of Ω_x, Ω_y, and Ωz, in which: $\Omega_x = cos(\theta) = \mu$; $\Omega_y = (1-\mu^2)^{\frac{1}{2}}\cos(\omega)$; and $\Omega_z = (1-\mu^2)^{\frac{1}{2}}\sin(\omega)$.

A Streamline Upwind Petrov-Galerkin (SUPG) formulation [58, 59] is used to spatially discretise the steady state Boltzmann transport equation. Prior to spatial discretisation, the angularly discretised Boltzmann transport equation is premultiplied by a spatial SUPG term to yield:

$$\left(\mathbf{I} - \mathbf{A}\cdot\nabla\mathbf{P}\right)\left(\mathbf{A}\cdot\nabla\mathbf{\Psi}(\mathbf{r}) + \mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r}) - \mathbf{S}(\mathbf{r})\right) = 0, \tag{5.10}$$

where $\mathbf{P}$ represents a stabilisation matrix and $\mathbf{H}(\mathbf{r}) = \mathbf{H}$. In the case of spherical harmonics, $\mathbf{P}$ is a matrix that determines the amount of weighting assigned to a given finite element discretisation. A matrix operator, $g(h\mathbf{\Lambda}_{H_{\nu\nu}})$, is used to simplify the description of stabilisation matrices. A new stabilisation matrix $\mathbf{P}$ is defined that is differentiable for inverse problems. The scattering/removal operator, $\mathbf{H}$, may be defined in terms of a matrix of right eigenvectors, $\mathbf{R}_H$, a matrix of left eigenvectors, $\mathbf{R}_H^{-1}$ and a corresponding diagonal matrix of eigenvalues of $\mathbf{\Lambda}_H$:

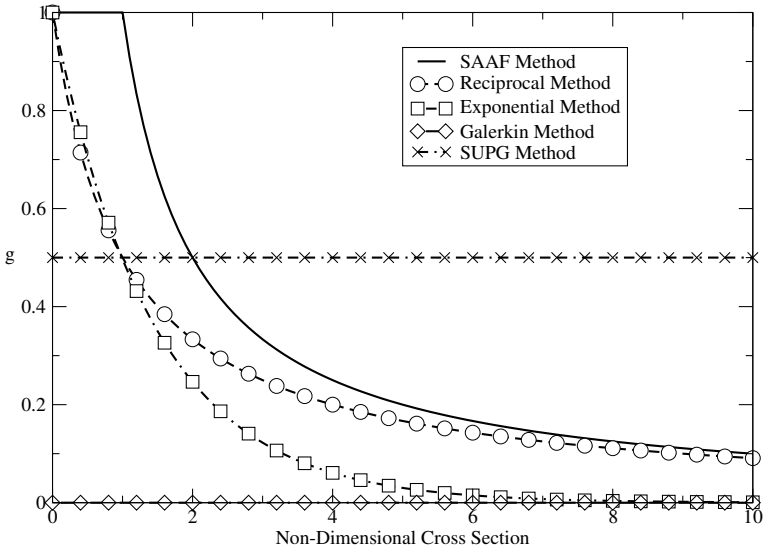


Figure 5.1: Examples of stabilisation functions (see equation 5.14) used in linear Petrov-Galerkin methods. Examples include stabilisation functions used in the SAAF method, the Reciprocal method, the Exponential method, the Galerkin method and the SUPG method.

$$\mathbf{H} = \mathbf{R}_H \mathbf{\Lambda}_H \mathbf{R}_H^{-1}. \tag{5.11}$$

In which $\mathbf{H}$ represents a positive, semi-definite matrix of non-negative eigenvalues. When using spherical harmonic basis functions $\mathbf{H}$ is a diagonal matrix and therefore $\mathbf{R}_H$ is the identity matrix. The matrix of modified eigenvalues is defined as:

$$\mathbf{G}_{\rho\nu}(\mathbf{\Lambda}_H) = \begin{cases} g(\mathbf{\Lambda}_{H_{\nu\nu}}) & if \qquad \rho = \nu \\ 0 & otherwise \end{cases} \tag{5.12}$$

The stabilisation matrix is subsequently defined:

$$\mathbf{P} = h\mathbf{R}_H \mathbf{G}(\mathbf{\Lambda}_H)\mathbf{R}_H^{-1}. \tag{5.13}$$

in which h is the spatially varying characteristic element size. Some specific examples of stablisation matrices used in linear Petrov-Galerkin techniques are described in figure 5.1. Examples include:

$$\begin{aligned}
\text{SAAF method,} \quad & g(h\mathbf{\Lambda}_{H_{\nu\nu}}) = \frac{1}{h\Lambda_{H_{\nu\nu}}}. \\
\text{Galerkin method,} \quad & g(h\mathbf{\Lambda}_{H_{\nu\nu}}) = 0. \\
\text{SUPG method,} \quad & g(h\mathbf{\Lambda}_{H_{\nu\nu}}) = \frac{1}{2}. \\
\text{Reciprocal method,} \quad & g(h\mathbf{\Lambda}_{H_{\nu\nu}}) = \frac{1}{1+h\mathbf{\Lambda}_{H_{\nu\nu}}}. \\
\text{The Exponential method,} \quad & g(h\mathbf{\Lambda}_{H_{\nu\nu}}) = e^{-0.37h\mathbf{\Lambda}_{H_{\nu\nu}}}.
\end{aligned}$$

All proceeding calculations have been performed using a $\mathbf{P}$ matrix designed by the exponential method above. The exponential $\mathbf{P}$ matrix provides an optimal method to use in both optically thick and optically thin media. If the second set of brackets on the left hand side of equation 5.10 is defined as the residual $\mathscr{R}(\mathbf{\Psi}(\mathbf{r}))$, equation 5.10 can be recast as:

$$(\mathbf{I} - \mathbf{A}\cdot\nabla\mathbf{P})\mathscr{R}(\mathbf{\Psi}(\mathbf{r})) = 0. \tag{5.14}$$

The spatial dependence of the angularly discretised flux, source and scattering/removal term are discretised in space using a finite element approximation [119]:

$$\begin{aligned}
\boldsymbol{\psi}(\mathbf{r}) \approx \mathbf{\Psi}(\mathbf{r}) &= \sum_{j=1}^{\mathcal{N}} \mathbf{N}_j(\mathbf{r})\mathbf{\Psi}_j, \\
\mathbf{s}(\mathbf{r}) \approx \mathbf{S}(\mathbf{r}) &= \sum_{j=1}^{\mathcal{N}_s} \mathbf{N}_{S_j}(\mathbf{r})\mathbf{S}_j, \\
\boldsymbol{\mathcal{H}}(\mathbf{r}) \approx \mathbf{H}(\mathbf{r}) &= \sum_{j=1}^{\mathcal{N}_H} \mathbf{N}_{H_j}(\mathbf{r})\mathbf{H}_j,
\end{aligned} \tag{5.15}$$

where $\mathbf{N}_j(\mathbf{r})$, $\mathbf{N}_{s_j}(\mathbf{r})$, and $\mathbf{N}_{H_j}(\mathbf{r})$ are each $(\mathcal{M}\times\mathcal{M})$ diagonal matrices containing finite element basis functions associated with the angular fluxes, sources and scattering/removal

terms respectively. $\mathcal{N}$, $\mathcal{N}_s$ and $\mathcal{N}_H$ represent the number of spatial finite element basis functions associated with the angular flux, sources and materials and $\mathbf{\Psi}_j$, $\mathbf{S}_j$, and $\mathbf{H}_j$ are vectors -of size $\mathcal{M}$- containing the moments of the angularly discretised flux, source, and scattering/removal terms. The spatially dependent finite element basis functions associated with the source and scattering/removal operator are spatially constant over a given element. Therefore both $\mathcal{N}_s$ and $\mathcal{N}_H$ represent the number of elements contained in the finite element mesh. To complete the discretisation, a Petrov-Galerkin method is used to reduce the Boltzmann transport equation into a series of ν linear equations. This is achieved by multiplying equation 5.10 by a further matrix of spatial finite element basis functions, $\mathbf{N}_i(\mathbf{r})$, before integrating the expression over the solution domain. Furthermore, Green's theorem [59] is used to convert the volume integrals representing the angularly discretised streaming terms into new volume integrals and surface terms.

$$
\begin{aligned}
- & \int_V \left(\mathbf{A}\cdot\nabla\mathbf{N}_i(\mathbf{r})\right)\mathbf{\Psi}(\mathbf{r})dV + \int_V \mathbf{N}_i(\mathbf{r})\left(\mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r}) - \mathbf{S}(\mathbf{r})\right)dV \\
+ & \int_V \mathbf{A}\cdot\nabla\mathbf{N}_i(\mathbf{r})\mathbf{P}\mathscr{R}dV + \int_\Gamma \mathbf{N}_i(\mathbf{r})\left(\mathbf{A}\cdot\mathbf{n}\right)\mathbf{\Psi}(\mathbf{r})d\Gamma \\
- & \int_\Gamma \mathbf{N}_i(\mathbf{r})\left(\mathbf{A}\cdot\mathbf{n}\right)\mathbf{P}\,\mathscr{R}d\Gamma = 0.
\end{aligned}
\tag{5.16}
$$

In which, V, represents the solution domain and, Γ, the surface of the domain. The finite element basis function $\mathbf{N}_i(\mathbf{r})$ is an $\mathcal{M}\times\mathcal{M}$ diagonal matrix containing the terms:

$$
\mathbf{N_i}(\mathbf{r}) = \begin{bmatrix}
N_i(\mathbf{r}) & 0 & . & . & 0 \\
0 & N_i(\mathbf{r}) & 0 & . & . \\
. & 0 & N_i(\mathbf{r}) & . & . \\
. & . & . & . & . \\
. & . & . & . & 0 \\
0 & . & . & 0 & N_i(\mathbf{r})
\end{bmatrix}, \tag{5.17}
$$

where $i = (1, 2, ..., \mathcal{N})$. A full description of the spatial discretisation of the Boltzmann

transport equation using Streamline Upwinded Petrov-Galerkin methods is described in [58, 59, 61]. [58, 59, 61] also describe how estimations of the angular flux at cell faces and domain boundaries are determined through a method of mapping the angular moments to Riemann variables. Following discretisation, the direct problem is defined using the matrix equation:

$$\mathbf{E}\boldsymbol{\Psi}_s = \mathbf{b_s}. \tag{5.18}$$

A total of $\mathcal{S}$ source problems must be solved. Each source problem results in a solution vector for the angular flux $\boldsymbol{\Psi}_s = (\boldsymbol{\Psi}_{\mathbf{s}\,1\,1}\ \boldsymbol{\Psi}_{\mathbf{s}\,1\,2}\ ...\ \boldsymbol{\Psi}_{\mathbf{s}\,\mathcal{NM}})^T$, in which $s = (1, 2, ..., \mathcal{S})$. By solving the discrete linear system described in equation 5.18, a solution for the angular flux at all points along the finite element mesh is found. The GMRES iterative solver is used to solve the discrete linear system [164]. The GMRES solver incorporates block forward and backward Gauss Siedel preconditioning and, for a large number of angular moments, an angular multigrid preconditioner. Each block is an $\mathcal{M} \times \mathcal{M}$ submatrix associated with each node of the finite element mesh. Matrix $\mathbf{E}$ depends on the discretisation method that is used and has a block sparse structure and vector, $\mathbf{b}_s = (\mathbf{b_s}_{1\,1}\ \mathbf{b_s}_{1\,2}\ ...\ \mathbf{b_s}_{\mathcal{NM}})^T$, is a boundary vector representing discretised source contributions. The finite element integrals used to calculate matrix $\mathbf{E}$ and source vector, $\mathbf{b_s}$, are evaluated numerically using Gauss quadrature. The localised finite-element basis functions used in the Petrov-Galerkin method are only non-zero in small regions of the domain, making the discretised system of equations sparse.

5.2.2 The Treatment of Sources and Detectors

In operator form, the forward model describes the angular flux, ψ_s, that results from a source placed in region, $\mathbf{r}$ at a boundary location, $\mathbf{r}_s$:

$$[\boldsymbol{\Omega} \cdot \boldsymbol{\nabla} + \mathcal{H}]\, \psi_s = C_s \delta(\mathbf{r} - \mathbf{r}_s). \tag{5.19}$$

Equation 5.19 describes a series of $\mathcal{S}$ source problems, where $s = (1, 2, ..., \mathcal{S})$. The source is a Dirac-delta function centred at position $\mathbf{r}_s$ of the domain $\mathbf{r}$. The magnitude of the source strength is given by C_s. ψ_s represents the angular flux owing to the source number, s, while the scattering/removal operator, $\mathcal{H}$, contains the spatially varying material properties. Each source problem yields an angular flux solution at all points in the solution domain. If the source position is located at a node then the source strength and angular flux are equal to the finite element basis function centred on that node multiplied by a scalar (to scale the strength):

$$\mathbf{b}_{s_i} = \int_V \mathbf{N}_i(\mathbf{r})\mathbf{S}_s dV, \tag{5.20}$$

in which $\mathbf{S}_s$ is a finite element approximation of the source term, $\mathbf{N}_i(\mathbf{r})$ represents a finite element basis function and $\mathbf{b}_s$ represents the magnitude of the source strength. If the source is located inside an element it must be distributed to the nearest node associated with that element and possibly nodes associated with neighbouring elements. A Gaussian function may be used to interpolate photons from a source on to adjacent nodes. If $\mathcal{N}$ represents the total number of nodes in the finite element mesh, the finite element approximation described in equations 5.5 and 5.15 is therefore defined according to the nodal values $\mathbf{S}_{si}$ of $\mathbf{S}_s$. Thus, for the i^{th} nodal position, $\mathbf{x}_i$, and the s^{th} source at position $\mathbf{r}_s$, $\mathbf{S}_{si}$ is defined at each node according to:

$$\mathbf{S}_{si} = \frac{1}{n_s} exp\left[-\frac{(\mathbf{x}_i - \mathbf{r}_s)^2}{l_s^2}\right]. \tag{5.21}$$

The angular flux received at a detector location is also calculated using a Gaussian function to interpolate the flux from any adjacent nodes on the surface of the detector. Detectors are distributed to the nearest nodes in a similar way as sources are distributed (i.e. according to equations 5.20, and 5.21). Thus, for the i^{th} nodal position, $\mathbf{x}_i$, the r^{th} detector at position $\mathbf{r}_s^r$, and the s^{th} source at position $\mathbf{r}_s$, the finite element approximation is defined at each node according to:

$$\mathbf{S}^r_{si} = \frac{1}{n^r_s} exp\left[-\frac{(\mathbf{x}_i - \mathbf{r}^r_s)^2}{l^2_r}\right]. \tag{5.22}$$

In equations 5.21, and 5.22, l_s, and l_r represent the width of the Gaussian, which controls the spreading of sources on to the nearest nodes and the spreading of angular flux from adjacent nodes on to the detectors. The normalisation coefficients, n_s or n^r_s are usually chosen to be equal to the integral value in equation 5.20 for source and detector contributions respectively. Unless a source or detector is placed at a node (in which case the Gaussian width is made very small) the Gaussian width is usually chosen to be the same size as the elements within the vicinity of the sources and detectors.

5.3 The Functional

The functional that is optimised to find a solution to the inverse problem is given by:

$$F = F_d + F_r + F_v. \tag{5.23}$$

In which F_d represents the data misfit between observed and predicted data and F_r and F_v are regularisation terms that encourages certain types of model. The model covariance information contained in the regularisation term applies smoothness and anisotropy constraints affecting a homogeneous isotropic model.

5.3.1 The Data Misfit

The functional is a quantification of the data misfit between the model predictions and measured data:

$$F_d = \frac{1}{2}\sum_{s=1}^{\mathcal{S}}\sum_{r=1}^{\mathcal{R}}\sum_{\mu=1}^{\mathcal{M}}\sum_{i=1}^{\mathcal{N}} W^r_{s\,i}\left(\Psi_{s\,\mu\,i} - d^r_{s\,\mu}\right)^2. \tag{5.24}$$

Equation 5.24 represents the weighted sum of the squared differences between observed data, $d^r_{s\ \mu}$, and data predicted from a given material property distribution, $\Psi_{s\ \mu\ i}$, for all, $\mathcal{S}$ sources and $\mathcal{R}$ detectors. The summation is also performed over each node, $i = (1, 2, ..., \mathcal{N})$, and for each angular moment $\mu = (1, 2, ..., \mathcal{M})$. The weighting factor, $W^r_{s\ i}$, distributes both the sources and detectors to the nearest nodes using a Gaussian approximation (see equation 5.28). In matrix form, the data misfit functional is given by:

$$F_d = \frac{1}{2}(\hat{\mathbf{\Psi}} - \mathbf{d})^{\mathbf{T}}\mathbf{W}(\hat{\mathbf{\Psi}} - \mathbf{d}), \tag{5.25}$$

in which

$$\hat{\mathbf{\Psi}}^T = (\mathbf{\Psi}_1^T\ \ \mathbf{\Psi}_2{}^T\ \ \cdots\ \ \mathbf{\Psi}_{\mathcal{S}}{}^T,\ \ \cdots\ \ , \mathbf{\Psi}_1{}^T\ \ \mathbf{\Psi}_2{}^T\ \ \cdots\ \ \mathbf{\Psi}_{\mathcal{S}}{}^T), \tag{5.26}$$

and

$$\mathbf{d}^T = (\mathbf{d}_1^{1\,T}\ \ \mathbf{d}_2^{1\,T}\ \ \cdots\ \ \mathbf{d}_{\mathcal{S}}^{1\,T},\ \ \cdots\ \ , \mathbf{d}_1^{\mathcal{R}\,T}\ \ \mathbf{d}_2^{\mathcal{R}\,T}\ \ \cdots\ \ \mathbf{d}_{\mathcal{S}}^{\mathcal{R}\,T}). \tag{5.27}$$

The data covariance matrix $\mathbf{W}^{-1}$ is diagonal and contains the weights, w^r_s associated with each datum. The vector, $\mathbf{\Psi} - \mathbf{d}$, represents the residuals (data misfit) at each node for a given source problem and for a given detector or detector pair.

5.3.2 The Gaussian Weight

For the i^{th} node at position $\mathbf{x}_i$, the r^{th} detector at position, $\mathbf{r}^r_s$, and the s^{th} source term, the weighting contribution is defined:

$$W^r_{s\ i} = \frac{w^r_s}{n^r_s} exp\left[-\frac{(\mathbf{x}_i - \mathbf{r}^r_s)^2}{l^2_r}\right]. \tag{5.28}$$

Where l_r is the length over which the interpolation is performed and n^r_s is a normalisation coefficient. Equation 5.28 provides a mechanism for gathering the flux from local nodes when the detector is placed between nodes. If the size of the Gaussian width, l_r, is small

compared to the element size then equation 5.28 assumes that the detector is re-positioned at the nearest node. In all examples in this thesis the detector positions are centred on nodes and l_r is small relative to the neighbouring element sizes. Thus the use of equation 5.28 takes the flux at the node in question and compares it with the idea detector response in equation 5.24. Therefore, $W_s^{r_k} = w_s^r/n_s^r$ for a single node $k = j$ (j is the node at which the detector is placed) otherwise, $W_s^{r_k} = 0$. The weighting function typically has three contributions. In the first two contributions w_s^r is used to weight the datum associated with each source-detector pair according to 1) the confidence in the datum and 2) for preferential weighting of datum. When using w_s^r for preferential weighting, large weights are typically assigned to small data-values. The final contribution in the weight, $W_{s\,i}^r$, arises from the way in which the detectors are treated. The detectors are distributed to the nearest node using a Gaussian distribution (equation 5.22) in the same way as sources are distributed. The weight function, w_s^r, is defined:

$$w_s^r = \begin{cases} 0 & if \quad ((x_r - x_s)^2 + (y_r - y_s)^2 + (z_r - z_s)^2) < \kappa^2 \\ \left(\dfrac{d_{max_s}}{\phi_1 d_{max_s} + \phi_2 d_{max_s}^r} \right)^2 & otherwise. \end{cases}$$

where $\kappa^2 = (l_s + l_r)^2$ and l_s and l_r represent the length over which the interpolation is performed. x_s, x_r, y_s, y_r, z_s and z_r represent the x, y and z position coordinates for the sources and detectors respectively. ϕ_1 and ϕ_2 are carefully chosen scalar constants that are used to adjust weights so that the data observed at detectors with relatively small scalar fluxes can be given higher priority. κ is chosen to avoid putting excessive emphasis on short range inversion information obtained by placing sources and detectors close to one another. d_{max_s} is associated with the maximum value of the observed angular flux for all detectors:

$$d_{max_s} = \max_r \{d_{max_s}^r\}, \tag{5.29}$$

and $d_{max_s}^r$ is the maximum value of the observed angular flux at detector, r, for source experiment, s.

5.3.3 Regularisation of Ill-Posed Inversion Problems

A problem is well-posed if its solution is unique and exists for arbitrary data. Most inversion problems are highly non-linear and ill-posed. That is, the solutions to inversion problems are generally either not unique, not continuous or do not exist for arbitrary data. Model covariance regularisation overcomes the problem of ill-posedness at the expense of limiting the allowed models to a class of models that are compatible with the provided model covariance information. Depth and directionally dependent model covariance matrices allow spatial variation and preferential model regularisation in arbitrary directions. Data weighting poorly resolved inversion regions can also be incorporated into regularisation matrices to yield homogeneous solutions. Using appropriate model covariance constraints and including any available a priori information can therefore remedy any concerns associated with non-uniqueness and limited data coverage.

Spatial Regularisation

Regularisation in medical imaging is designed so that the layering of biological tissue may be taken into account. In layered tissue environments the subsurface is smooth in the layering plane but rough perpendicular to the layering plane. Spatial regularisation is used to constrain the anisotropic and inhomogeneous structural properties of tissue, and to incorporate these properties into the inversion along with any useful a priori information. The regularisation term defined in equation 5.23 is defined as:

$$F_r = \frac{1}{2}\gamma_a \int_V \nabla^T m_a k \, \nabla m_a \, dV + \frac{1}{2}\gamma_{s0} \int_V \nabla^T m_{s0} k \, \nabla m_{s0} \, dV = \frac{1}{2}\mathbf{m}^T \mathbf{K} \mathbf{m}, \quad (5.30)$$

$\mathbf{K}$ represents a diffusion tensor and is assumed to be isotropic and equal to one. γ_a and γ_{s0} control the level of model covariance imposed for absorption and scattering contributions respectively. $\mathbf{m}_a = (m_{a1}\ m_{a2}\ ...\ m_{a\mathcal{N}_H})^T$ represents the discretised absorption material properties where -assuming m_{a_i} is the nodal value of $m_{a_i} = ln(\mu_{a_i})$. $\mathbf{m}_{s0} = (m_{s01}\ m_{s02}\ ...\ m_{s0\mathcal{N}_H})^T$ represents the discretised scattering material properties in

which $m_{s0_i} = ln(\mu_{s0_i})$. $\mathbf{m}_a$ and $\mathbf{m}_{s0}$ are defined collectively in terms of $\mathbf{m}^T = (\mathbf{m}_a^T\ \mathbf{m}_{s0}^T)$, and are approximated using a set of finite element basis functions, N_{H_j}, such that:

$$m_a = \sum_{j=1}^{\mathcal{N}_H} N_{H_j} m_{a_j}, \qquad m_{s0} = \sum_{j=1}^{\mathcal{N}_H} N_{H_j} m_{s0_j}, \tag{5.31}$$

Material properties are considered as opposed to material cross-sections because the resulting contributions to the error functional are quadratic. The error functional is therefore easily optimised using standard least squares techniques, resulting in uniform material properties. In addition, material cross-sections, sometimes vary by several orders of magnitude, which is undesirable when considering model regularisation. Structure is penalised using a symmetric, positive-definite matrix, $\mathbf{K}$:

$$\mathbf{K}_{ij} = \begin{bmatrix} \gamma_a \int_V \nabla N_{H_i}\ k\ \nabla N_{H_j}\ dV & 0 \\ 0 & \gamma_{s0} \int_V \nabla N_{H_i}\ k\ \nabla N_{H_j}\ dV \end{bmatrix}, \tag{5.32}$$

In which $\mathbf{K}$ represents a diffusion matrix and is defficient in rank by 1. Therefore, $\mathbf{K}$ is not invertible and a singular value has to be set in order to make $\mathbf{K}$ non-singular. Although $\mathbf{K}$ is not invertible, a combination of likelihood and prior information can be used to eliminate the problem null space so that a unique solution may be obtained. k is a positive semi-definite 3×3 matrix and defines a scalar product, making the integrand positive throughout the domain and therefore F_r positive. For medical imaging applications inhomogeneous smoothing is required, therefore each of the elements contained inside matrix $\mathbf{K}$ should be defined as a function of position. Large initial values of γ_a or γ_{s0} affect a smooth, homogeneous model allowing large scale structures to emerge during early iterations. During later iterations, structurally complex models emerge as, γ_a and γ_{s0}, are relaxed.

Penalising Deviation from the Starting Model

The penalty deviation associated with an updated model relative to the starting model is achieved by replacing the absorption material property, $\mathbf{m}_a$, with $\Delta\mathbf{m}_a = \mathbf{m}_a - \mathbf{m}_{a_{obs}}$

and the scattering material property, $\mathbf{m}_{s0}$, with $\Delta\mathbf{m}_{s0} = \mathbf{m}_{s0} - \mathbf{m}_{s0_{obs}}$:

$$\begin{aligned} F_v &= \frac{1}{2}\lambda \int_V (m_a - m_{a_{obs}})^2 dV + \frac{1}{2}\lambda \int_V (m_{s0} - m_{s0_{obs}})^2 dV \\ &\approx \frac{1}{2}\Delta\mathbf{m}_a^T \mathbf{V}\mathbf{M}_{H_L}\Delta\mathbf{m}_a + \frac{1}{2}\Delta\mathbf{m}_{s0}^T \mathbf{V}\mathbf{M}_{H_L}\Delta\mathbf{m}_{s0}, \end{aligned} \tag{5.33}$$

in which $\Delta\mathbf{m}_a = (\Delta m_{a1}\ \Delta m_{a2}\ ...\ \Delta m_{a\mathcal{N}_H})^T$ and $\Delta\mathbf{m}_{s0} = (\Delta m_{s01}\ \Delta m_{s02}\ ...\ \Delta m_{s0\mathcal{N}_H})^T$ represent the change in the absorption and scattering material properties. The absorption and scattering material properties associated with the previous least squares iteration are given by $\mathbf{m}_{a_{obs}} = (m_{a_{obs}1}\, m_{a_{obs}2}\, ..\, m_{a_{obs}\mathcal{N}_H})^T$ and $\mathbf{m}_{s0_{obs}} = (m_{s0_{obs}1}\, m_{s0_{obs}2}\, ..\, m_{s0_{obs}\mathcal{N}_H})^T$ respectively. $\mathbf{M}_{H_L}$ is a lumped diagonal mass matrix containing only diagonal non-zero values. In other words:

$$\mathbf{M}_{H_{L_{ij}}} = \begin{cases} M_{H_{L_{ij}}} \text{ if } i = j; \\ 0 \quad \text{otherwise.} \end{cases} \tag{5.34}$$

The lumped mass matrix, $\mathbf{M}_{H_L}$ is found using finite element basis functions N_{H_i}

$$\mathbf{M}_{H_{L_{ii}}} = \int_V N_{H_i} dV, \tag{5.35}$$

The Bayesian framework for regularisation interprets $\mathbf{M}_{H_L}$ as a priori information about the underlying image probability distribution. In other words, $\mathbf{M}_{H_L}$ is a penalty term. The penalty term discourages reconstructed images with high values of $\Delta\mathbf{m}_a^T \mathbf{V}\mathbf{M}_{H_L}\Delta\mathbf{m}_a$ or $\Delta\mathbf{m}_{s0}^T \mathbf{V}\mathbf{M}_{H_L}\Delta\mathbf{m}_{s0}$. However, in practice $\mathbf{M}_{H_L}$ is selected using ad hoc or heuristic methods. If the prior does not model the actual data, the reconstruction is biased [169]. The scalar λ controls the magnitude of the step length damping. Matrix $\mathbf{V}$ is defined:

$$\mathbf{V}_{ij} = \begin{cases} \lambda \text{ if } i = j; \\ 0 \text{ otherwise.} \end{cases} \tag{5.36}$$

5.4 Least Squares Inversion

The adjoint model [79, 162, 163] is a tool that has been developed for the inverse modelling of physical systems. The adjoint problem is based on the forward problem and a discrepancy (misfit) calculation between forward model predictions and the measured data. The discrepancy calculation is contained in a functional which is minimised using a gradient based optimisation technique [71, 72]. For medical imaging applications the systems control variables represent the material properties associated with the medium. Providing the initial model is close enough to the global solution to allow successive relinearisations, the model trajectory can be adjusted as close as possible to the measured data. The adjoint form of equation 5.24 is defined:

$$\mathbf{E}^T \mathbf{\Psi}_s^* = \frac{\partial F}{\partial \mathbf{\Psi}_s}, \tag{5.37}$$

where, $\mathbf{\Psi}_\mathbf{s}^* = (\mathbf{\Psi}_{\mathbf{s}\,1\,1}^* \; \mathbf{\Psi}_{\mathbf{s}\,1\,2}^* ... \mathbf{\Psi}_{\mathbf{s}\,\mathcal{NM}}^*)^T$ is the vector of the (adjoint) control variables and F represents the error functional defined in equation 5.23. Equation 5.37 describes a series of $\mathcal{S}$ source problems, where $s = (1, 2, ..., \mathcal{S})$. The differential term on the right hand side of equation 5.37 represents the data residual, which is used as a source term for adjoint calculations. Using the abbreviation $\boldsymbol{\epsilon} = \mathbf{\Psi} - \mathbf{d}$ to represent the data residual, the residual for an updated model may be approximated according to:

$$\boldsymbol{\epsilon}(m_{new}) = \boldsymbol{\epsilon}(m_{old}) + \mathbf{J}(\mathbf{m}_{new} - \mathbf{m}_{old}). \tag{5.38}$$

$\mathbf{m}_{old}$ represents the previously available best fit material properties and $\mathbf{m}_{new}$ denotes the updated material properties. The notation $\boldsymbol{\epsilon}(m_{old})$ and $\boldsymbol{\epsilon}(m_{new})$ indicates that the residual is a function of the old material properties and new material properties respectively. $\mathbf{J}$ is a Jacobian matrix:

$$\mathbf{J} = \frac{\partial \boldsymbol{\epsilon}(m_{old})}{\partial \mathbf{m}_{old}}. \tag{5.39}$$

The elements of the Jacobian matrix consists of the changes in the data misfit, $\boldsymbol{\epsilon}(m_{old})$,

for the source-detector pair, s and r, caused by a perturbation in the material properties $\mathbf{m}_{old}$. Using the approximation for the residuals given in equation 5.38 the error functional defined in equations 5.23, 5.30, and 5.33 becomes:

$$\begin{aligned} F &= \frac{1}{2}\left[\boldsymbol{\epsilon}(m_{old}) + \mathbf{J}(\mathbf{m}_{new} - \mathbf{m}_{old})\right]\mathbf{W}\left[\boldsymbol{\epsilon}(m_{old}) + \mathbf{J}(\mathbf{m}_{new} - \mathbf{m}_{old})\right] \\ &+ \frac{1}{2}\mathbf{m}_{new}^T\mathbf{K}\mathbf{m}_{new} \\ &+ \frac{1}{2}\lambda(\mathbf{m}_{new} - \mathbf{m}_{old})\mathbf{M}_{H_L}(\mathbf{m}_{new} - \mathbf{m}_{old}). \end{aligned} \tag{5.40}$$

Differentiating the functional (equation 5.40) with respect to the control variables, $\mathbf{m}_{old}$, and equating to zero yields the solution minimum:

$$(\mathbf{J}^T\mathbf{W}\mathbf{J} + \mathbf{K} + \lambda\mathbf{M}_{H_L})\Delta\mathbf{m} = -\mathbf{J}^T\mathbf{W}^T\boldsymbol{\epsilon}(m_{old}) - \mathbf{K}\mathbf{m}_{old}. \tag{5.41}$$

This least squares system for model updates $\Delta\mathbf{m} = \mathbf{m}_{new} - \mathbf{m}_{old}$ is called the Levenberg-Marquardt method. Once model updates are calculated the old model is then updated. If the inverse problem is non-linear, one can relinearise about the updated model and calculate a further model update. Equation 5.41 can be expressed:

$$\mathbf{B}\Delta\mathbf{m} = -\mathbf{a}, \tag{5.42}$$

where

$$\mathbf{B} = (\mathbf{J}^T\mathbf{W}\mathbf{J} + \mathbf{K} + \lambda\mathbf{M}_{H_L}), \tag{5.43}$$

and

$$\mathbf{a} = -\mathbf{J}^T\mathbf{W}^T\boldsymbol{\epsilon}(m_{old}) - \mathbf{K}\mathbf{m}_{old}. \tag{5.44}$$

The term $\lambda \mathbf{M}_{H_L}$ in equation 5.43 controls the magnitude of the model update $\Delta \mathbf{m}$. If λ is chosen to be large then $\Delta \mathbf{m}$ stays small. $\lambda \mathbf{M}_{H_L}$ is also used to increase the diagonal dominance and therefore the conditioning of matrix $\mathbf{B}$ in equation 5.42. $\mathbf{J}^T\mathbf{W}\mathbf{J}$ represents the approximate Hessian matrix and $\mathbf{a}$ represents the gradient.

5.4.1 Calculations Involving the Jacobian and Hessian Matrices

Both the Jacobian matrix, $\mathbf{J}$, and the approximate Hessian matrix, $\mathbf{J}^T\mathbf{W}\mathbf{J}$, are never explicitly computed or stored. Only the matrix-vector multiplication using these two matrices is required. The Jacobian matrix, $\mathbf{J}$, that is illustrated in equation 5.41 is defined:

$$\mathbf{J} = \mathcal{I}^T \begin{pmatrix} \mathcal{J}_1 \\ \mathcal{J}_2 \\ .. \\ .. \\ .. \\ \mathcal{J}_{\mathcal{S}} \end{pmatrix}, \tag{5.45}$$

in which $\mathcal{I}^T = (I_1\ I_2\ ...\ I_{\mathcal{S}})$ where each I_s for $s = (1, 2, ..., \mathcal{S})$ represents an identity matrix of order $\mathcal{N}_H \times \mathcal{N}_H$. Each $\mathcal{J}_s$ contains the change in the angular flux at a given node caused by a perturbation in the material properties at another node. The change in the angular flux with respect to the material property perturbations are calculated for each source forward problem, s. Multiplying the system by the $\mathcal{N}_H \times \mathcal{N}_H$ identity matrix $\mathcal{I}^T$ affects a summation of the individual Jacobians, $\mathcal{J}_s$ of every source experiment ($\mathcal{I}^T = (I_1\ I_2\ ...\ I_{\mathcal{S}})$ for $s = (1, 2, ..., \mathcal{S})$) into the Jacobian $\mathbf{J}$ for the entire data set. The Jacobian, $\mathcal{J}_s$ of the s^{th}-source experiment is approximated according to:

$$\mathcal{J}_s = \mathbf{E}^{-1}\mathbf{L}_s, \tag{5.46}$$

in which $\mathbf{L}_s = (\mathbf{l}_{a_{s1}}\ \mathbf{l}_{a_{s2}}\ ..\ \mathbf{l}_{a_{s\mathcal{N}_H}}\ \mathbf{l}_{s0_{s1}}\ \mathbf{l}_{s0_{s2}}\ ...\ \mathbf{l}_{s0_{\mathcal{S}\mathcal{N}_H}})$ and matrix $\mathbf{E}$ describes the direct

problem in equation 5.18. The j^{th} column vector, associated with absorption material properties, $\mathbf{l}_{a_{sj}}$ is defined as:

$$\mathbf{l}_{a_{sj}} = \frac{\partial \mu_{a_j}}{\partial m_{a_j}} \left(\frac{\partial \mathbf{E}}{\partial \mu_{a_j}} \mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial \mu_{a_j}} \right). \tag{5.47}$$

The j^{th} column vector, associated with scattering material properties, $\mathbf{l}_{s0_{sj}}$ is defined as:

$$\mathbf{l}_{s0_{sj}} = \frac{\partial \mu_{s0_j}}{\partial m_{s0_j}} \left(\frac{\partial \mathbf{E}}{\partial \mu_{s0_j}} \mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial \mu_{s0_j}} \right). \tag{5.48}$$

Both an absorption coefficient and a scattering coefficient are defined for each cell. The partial derivatives of the scattering and absorption coefficients with respect to material property control variables, m_{a_j} and m_{s0_j} have been defined $\frac{\partial \mu_{a_j}}{\partial m_{a_j}} = exp(m_{a_j})$ and $\frac{\partial \mu_{s0_j}}{\partial m_{s0_j}} = exp(m_{s0_j})$ respectively, which are exact for each node and ensures positivity. Furthermore the scattering/removal operator is a sparse diagonal matrix:

$$\mathbf{H} = \begin{bmatrix} \mu_a + \mu_{s0} - \mu_{s0} & 0 & . & . & 0 \\ 0 & \mu_a + \mu_{s0} - \mu_{s1} & 0 & . & . \\ . & 0 & \mu_a + \mu_{s0} - \mu_{s2} & 0 & . \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & \mu_a + \mu_{s0} - \mu_{s\mathcal{M}-1} \end{bmatrix} \tag{5.49}$$

The term contained in brackets in equation 5.47 can be formulated in terms of the gradient of the absorption material properties:

$$\frac{\partial \mathbf{E}}{\partial \mu_{a_j}}\boldsymbol{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial \mu_{a_j}} = \int_V \Big\{(\mathbf{N}_j - (\nabla \mathbf{N}_j \cdot \mathbf{A})\mathbf{P})\frac{\partial \mathbf{H}}{\partial \mu_{a_j}}\boldsymbol{\Psi}_s - (\nabla \mathbf{N}_j \cdot \mathbf{A})\left(\frac{\partial \mathbf{P}}{\partial \mu_{a_j}}\right)(\mathbf{A}\cdot\nabla\boldsymbol{\Psi}_s + \mathbf{H}\boldsymbol{\Psi}_s - \mathbf{S}_s)\Big\}dV, \quad (5.50)$$

and the gradient of the scattering material properties respectively:

$$\frac{\partial \mathbf{E}}{\partial \mu_{s0_j}}\boldsymbol{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial \mu_{s0_j}} = \int_V \Big\{(\mathbf{N}_j - (\nabla \mathbf{N}_j \cdot \mathbf{A})\mathbf{P})\frac{\partial \mathbf{H}}{\partial \mu_{s0_j}}\boldsymbol{\Psi}_s - (\nabla \mathbf{N}_j \cdot \mathbf{A})\left(\frac{\partial \mathbf{P}}{\partial \mu_{s0_j}}\right)(\mathbf{A}\cdot\nabla\boldsymbol{\Psi}_s + \mathbf{H}\boldsymbol{\Psi}_s - \mathbf{S}_s)\Big\}dV, \quad (5.51)$$

where $\mathbf{S}_s$ represent the source for source problem s. The scattering material properties incorporate both the scattering coefficient, μ_s, and the scattering moments, μ_{sl}. Matrix-vector multiplications involving, $\mathbf{L}_s$, or its transpose, $\mathbf{L}_s^T$, are demonstrated using the vectors, $\mathbf{q} = (q_{a_1}\ q_{a_2}\ ...\ q_{a_{\mathcal{N}_H}}\ q_{s0_1}\ q_{s0_2}\ ...\ q_{s0_{\mathcal{N}_H}})^T$ and $\mathbf{v} = (v_{a_1}\ v_{a_2}\ ...\ v_{a_{\mathcal{N}_H}}\ v_{s0_1}\ v_{s0_2}\ ...\ v_{s0_{\mathcal{N}_H}})^T$ respectively. For each length, $\mathcal{N}$, the matrix-vector multiplications take on the form

$$\mathbf{L}_s\mathbf{q} = (q_{a_1}\mathbf{l}_{a_{s1}} + q_{a_2}\mathbf{l}_{a_{s2}} + ..., q_{a_{\mathcal{N}_H}}\mathbf{l}_{a_{s\mathcal{N}_H}} + q_{s0_1}\mathbf{l}_{s0_{s1}} + q_{s0_2}\mathbf{l}_{s0_{s2}} + ..., q_{s0_{\mathcal{N}_H}}\mathbf{l}_{s0_{\mathcal{SN}_H}}), \quad (5.52)$$

$$\mathbf{L}_s^T\mathbf{v} = (v_{a_1}\mathbf{l}_{a_{s1}} + v_{a_2}\mathbf{l}_{a_{s2}} + ..., v_{a_{\mathcal{N}_H}}\mathbf{l}_{a_{s\mathcal{N}_H}} + v_{s0_1}\mathbf{l}_{s0_{s1}} + v_{s0_2}\mathbf{l}_{s0_{s2}} + ..., v_{s0_{\mathcal{N}_H}}\mathbf{l}_{s0_{\mathcal{SN}_H}})^T. \quad (5.53)$$

Since the nodal value for μ_{a_j} or μ_{s0_j} only manifests itself in $\mathbf{E}$ in a small region of the local vicinity of node, j, $\mathbf{l}_{a_{sj}}$ and $\mathbf{l}_{s0_{sj}}$ have very few non-zero entries. Manipulations that involve $\mathbf{l}_{a_{sj}}$ or $\mathbf{l}_{s0_{sj}}$ are therefore computationally efficient. First order methods are necessary conditions for optimality. To obtain a more accurate model the second order adjoint equation must be solved. The second order adjoint utilises the positive, definite

approximate Hessian matrix, $\mathbf{J}^T\mathbf{W}\mathbf{J}$, in addition to the first order gradient vector. The approximate Hessian, $\mathbf{J}^T\mathbf{W}\mathbf{J}$, is defined in terms of the matrix-vector product involving $\mathbf{L}_s$ and $\mathbf{L}_s^T$ with an arbitrary vector, $\mathbf{q}$, of length, $2\mathcal{N}_H$.

$$\mathbf{J}^T\mathbf{W}\mathbf{J}\mathbf{q} = \sum_{s=1}^{\mathcal{S}} \mathbf{L}_s^T\mathbf{E}^{-T}\mathcal{W}_s\mathbf{E}^{-1}\mathbf{L}_s\mathbf{q}, \tag{5.54}$$

where $\mathcal{W}_s = \sum_{r=1}^{\mathcal{R}} \mathbf{W}_s^r$ and the diagonal matrix $\mathbf{W}_s^r$ has diagonal elements w_s^r. The matrix $\mathbf{E}^{-1}$ defines a direct calculation while $\mathbf{E}^{-T}$ is the adjoint model. To obtain vector, $\mathbf{J}^T\mathbf{W}\mathbf{J}\mathbf{q}$, the solution of two matrix equations involving matrix $\mathbf{E}$ are required for each source problem. If the Hessian matrix was used instead of an approximate Hessian matrix, a further two matrix equations would have to solved for each source problem. It is unclear whether using the exact Hessian, instead of this approximation to it, would result in a more efficient algorithm.

5.5 Numerical Solution of the Inversion Problem

The functional F defined in equation 5.23 is minimised using a Levenberg-Marquardt optimisation algorithm. After calculating the model update, $\Delta\mathbf{m}$, the previous model is updated and the error functional is evaluated again using the updated model. Both the magnitude of the step length damping, λ and the regularisation (contained in the multipliers, γ_a and γ_{s0}) may be adjusted between the solution of the least squares problem.

5.5.1 The Solution Method

An iterative preconditioned linear conjugate gradient (PCG) solver is used to solve both the forward problem and the least-squares inverse problem. The equations that describe the PCG solver are described in detail by Pain et al [74]. The PCG solver requires the formation of a preconditioning matrix, $\hat{\mathbf{B}}$. The preconditioner is usually defined in such a way that the system, $\hat{\mathbf{B}}\Delta\mathbf{m} = \mathbf{a}$, amounts to a simplified version of the original system,

$\mathbf{B}\Delta\mathbf{m} = \mathbf{a}$ (defined in equation 5.43). The preconditioner, $\hat{\mathbf{B}}$, therefore represents an approximation to the approximate Hessian matrix, $\mathbf{B}$ and is defined:

$$\hat{\mathbf{B}} = \mathbf{K} + \lambda \mathbf{M}_{H_L}. \tag{5.55}$$

The preconditioner consists of the matrices containing structural constraints and step length damping. This enables model covariance to be rigorously enforced during each PCG iteration. The matrix, $\hat{\mathbf{B}}$, is non-singular and symmetric positive definite for a positive non-zero λ. Convergence to a local minimum is assumed when the maximum percentage difference in the material cross-section updates from two consecutive iterations becomes less than a certain tolerance.

Choice of Step length Damping Coefficient

The step length damping coefficient, λ (described in equation 5.33), controls the convergence of the PCG solution in two ways. Firstly λ controls the conditioning (i.e. the diagonal dominance) of matrix $\mathbf{B}$ to ensure that $\mathbf{B}$ is not singular. Secondly, the preconditioner, $\hat{\mathbf{B}}$, contains the diagonal matrix $\lambda\mathbf{M}_{H_L}$, which is also contained in $\mathbf{B}$. Increasing λ thus makes $\hat{\mathbf{B}}$ a better approximation to $\mathbf{B}$. A suitably large starting value of λ is given and the level of step length damping λ is adjusted according to the following heuristics: 1) λ is increased by a factor of 10 if either the PCG method is not converging well or the updated model is worse than the previous model; 2) If using a larger step length damping coefficient does not decrease the functional, F, the previous step length damping coefficient is used. 3) If PCG solution of $\mathbf{B}\Delta\mathbf{m} = \mathbf{a}$ is converging well [74] and the new functional is less than the old functional, then λ is reduced by a factor of 10.

Choice of Penalty Level

The structure penalty levels, γ_a and γ_{s0} (described in equation 5.30), influence the relative importance of the structure penalty to the error functional for absorption and scattering contributions respectively. γ_a and γ_{s0} are defined:

$$\gamma_{a_l} = \nu_a \gamma_{a_{l-1}}, \qquad \gamma_{s0_l} = \nu_{s0} \gamma_{s0_{l-1}}, \tag{5.56}$$

where l represents the iteration level. The relaxation factors ν_a and ν_{s0} need to be chosen to be less than or equal to one. The strategy for choosing the penalty levels, γ_a and γ_{s0}, is to initially use a large regularisation penalty level and gradually, at each subsequent Levenberg-Marquardt step, relax the respective penalty levels using ν_a and ν_{s0}. This forces large-scale structure to emerge during early iterations. During later iterations structurally more complex models emerge. The inversion process is stopped when the observed data is predicted to a pre-determined level or the model-updates become negligibly small. At that point a model is found that is as smooth as possible and as rough as necessary to fit the data to a sufficient level. Experience suggests the differentiation involved in the least squares method must be exact. The differentiation must be performed at a discrete level to provide gradients consistent with the data-misfit functional and the discretised equations.

5.6 Numerical Examples

Presented are the simultaneous reconstructions of absorption and scattering coefficients contained inside a given isotropic medium from measurements taken at the surface of the domain. Circular and square test objects containing embedded inhomogeneities with different scattering and absorption parameters are utilised as a crude representation of the human head or breast. The forward (light transport) model used in the reconstruction of each of the test cases was simulated using the steady state finite element representation of the time independent Boltzmann transport equation described in the theory. The inversion calculations utilise an iterative, adjoint reconstruction scheme that repeatedly calculates an update of the solution until an error functional, becomes acceptably small. The various locations of the individual sources and detectors for each of the subsequent experiments are described in tables 5.1 and 5.2.

	Experiment 1		Experiment 2		Experiments 3 and 4	
Source Number	x	y	x	y	x	y
1	-2.5	0.0	0.0	3.5	1.0	0.0
2	-1.75	1.75	3.5	0.0	2.0	0.0
3	0.0	2.5	0.0	-3.5	3.0	0.0
4	1.75	1.75	-3.5	0.0	4.0	0.0
5	2.5	0.0			5.0	1.0
6	1.75	-1.75			5.0	2.0
7	0.0	-2.5			5.0	3.0
8	-1.75	-1.75			5.0	4.0
9					4.0	5.0
10					3.0	5.0
11					2.0	5.0
12					1.0	5.0
13					0.0	4.0
14					0.0	3.0
15					0.0	2.0
16					0.0	1.0

Table 5.1: The x co-ordinate and y co-ordinate source locations for each of the steady state inversion experiments described in chapter 5.

	Experiment 1		Experiment 2		Experiments 3 and 4	
Detector Number	x	y	x	y	x	y
1	-2.5	0.0	0.0	3.5	1.0	0.0
2	-1.75	1.75	2.5	2.5	2.0	0.0
3	0.0	2.5	3.5	0.0	3.0	0.0
4	1.75	1.75	2.5	-2.5	4.0	0.0
5	2.5	0.0	0.0	-3.5	5.0	1.0
6	1.75	-1.75	-2.5	-2.5	5.0	2.0
7	0.0	-2.5	-3.5	0.0	5.0	3.0
8	-1.75	-1.75	-2.5	2.5	5.0	4.0
9					4.0	5.0
10					3.0	5.0
11					2.0	5.0
12					1.0	5.0
13					0.0	4.0
14					0.0	3.0
15					0.0	2.0
16					0.0	1.0

Table 5.2: The x co-ordinate and y co-ordinate detector locations for each of the steady state inversion experiments described in chapter 5.

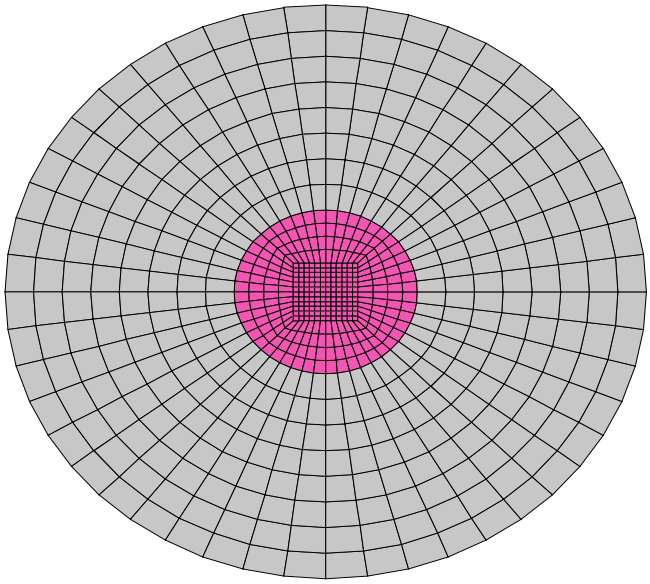


Figure 5.2: A two dimensional structured finite element circular mesh of diameter 5cm containing an inhomogeneity of diameter 2cm located at the centre of the domain. The steady state problem has 1186 nodes and 1760 bi-linear quadrilateral elements.

5.6.1 Experiment 1

The first simulation geometry depicts a circle 5cm in diameter containing an inhomogeneity situated in the centre of the circle that is 2cm in diameter (figure 5.2). The absorption and scattering coefficients of the medium and the inhomogeneity are contained in table 5.3. Eight equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward model and reconstruction model results. P_1 and P_3 angular approximations are used to obtain the steady state forward and inversion calculations along with a Levenberg-Marquardt optimisation method. The same sized mesh is used to generate both the synthetic data and the inversion reconstruction. The mesh used is two dimensional and has 1186 nodes and 1760 bi-linear quadrilateral elements. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium was chosen with optical properties equal to those of the background medium (table 5.3). In other words $\mu_a = 0.2$cm and $\mu_s^{'} = 7.0$cm.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.2	7.0
Inhomogenity	0.1	5.0

Table 5.3: Optical absorption and scattering coefficients for a circle 5cm in diameter containing an inhomogeneity placed at the centre of that circle 2cm in diameter.

Figure 5.3 illustrates scalar flux (i.e. the zeroth moment angular flux) as a function of x- and y- spatial co-ordinates for the forward solution (top left hand side graph), the natural log forward solution (top right hand side graph) and the adjoint solution (bottom graph) for a single source-detector arrangement. Figure 5.3 illustrates the results obtained using a P_3 angular expansion. The peak's illustrated in the forward solution diagrams indicate the position of the isotropic source (at co-ordinates -2.5, 0.0) and how the scalar flux propagates through the medium. The adjoint solution represents the data misfit between the exact solution and the model prediction. The adjoint solution illustrates the adjoint flux obtained from a detector placed at co-ordinates (-1.75, 1.75). The first and second order gradient of the error functional attempt to adjust the absorption and scattering material properties so that the forward model can better fit the exact solution obtained at the detectors. As a result the scalar flux of the adjoint solution, the error functional, and the gradient of the error functional all tend to zero as a more accurate angular flux approximation is found. The speed of convergence depends on the optimisation technique used, the number of iterations performed, the refinement of the finite element mesh and the complexity of the problem.

Figure 5.4 compares the converged absorption (left hand side) and scattering (right hand side) reconstruction results for the problem geometry described in figure 5.2 and table 5.3. P_1 (top) and P_3 (bottom) angular approximations were used to perform the inversion calculations. The results confirm that a P_1 angular approximation provides an adequate description for modelling radiation transport in highly scattering media. Figure 5.4 provides visual evidence that during P_1 and P_3 inversion calculations, the algorithm finds the correct location and approximate size of the inhomogeneity for both the absorption and scattering reconstructions. The relative magnitudes of the absorption and scattering

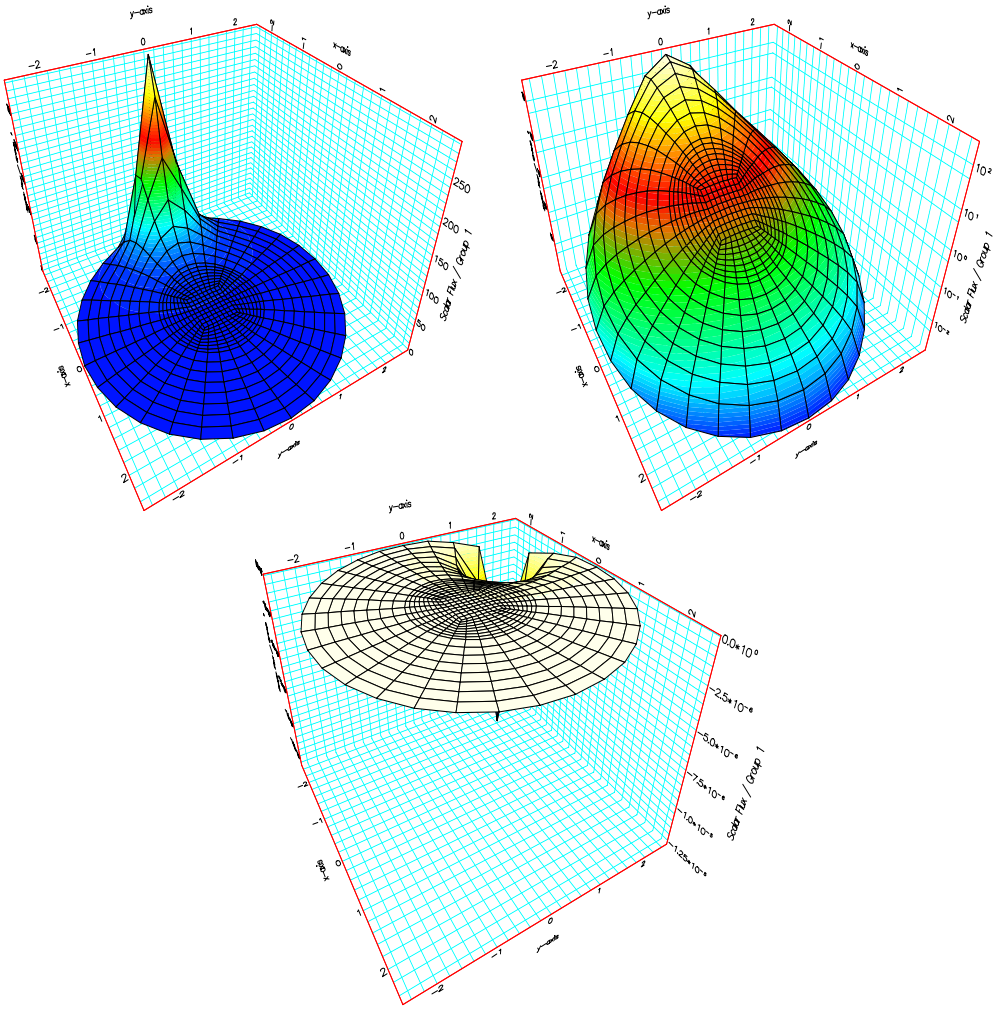


Figure 5.3: Scalar flux as a function of two dimensional space (x and y) for the forward solution (top left hand side graph), the natural log of the forward solution (top right hand side graph) and the adjoint solution (bottom graph). A P_3 angular approximation is used to obtain each of the above images. Scalar flux represents the zeroth moment angular flux.

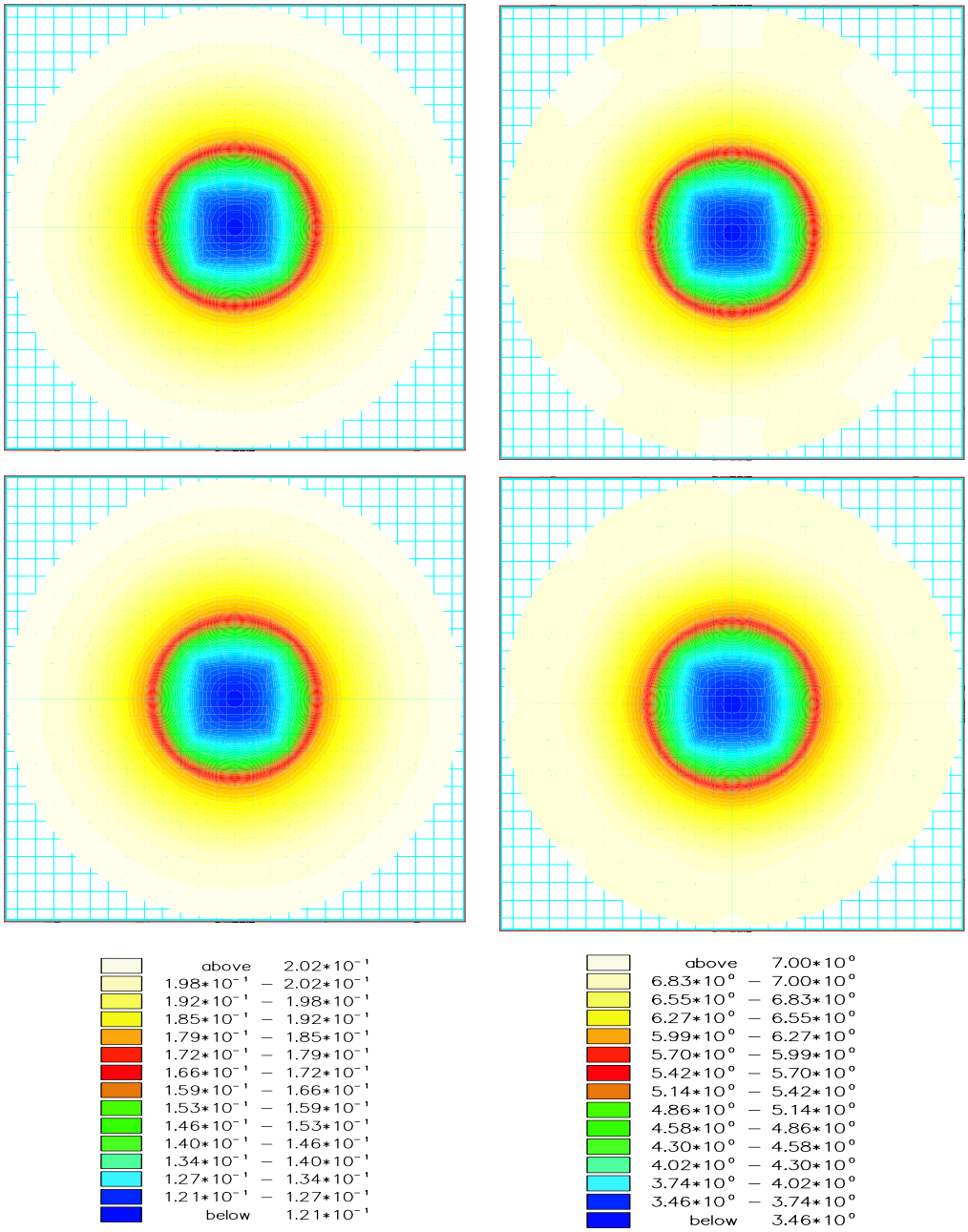


Figure 5.4: Absorption (left hand side) and scattering (right hand side) reconstruction results for a circle in circle type problem using P_1 (top) and P_3 (bottom) angular approximations. The results confirm that a P_1 angular approximation provides an adequate description for modelling radiation transport in highly scattering media. For each angular approximation, convergence was reached after 11 Levenberg-Marquardt iterations.

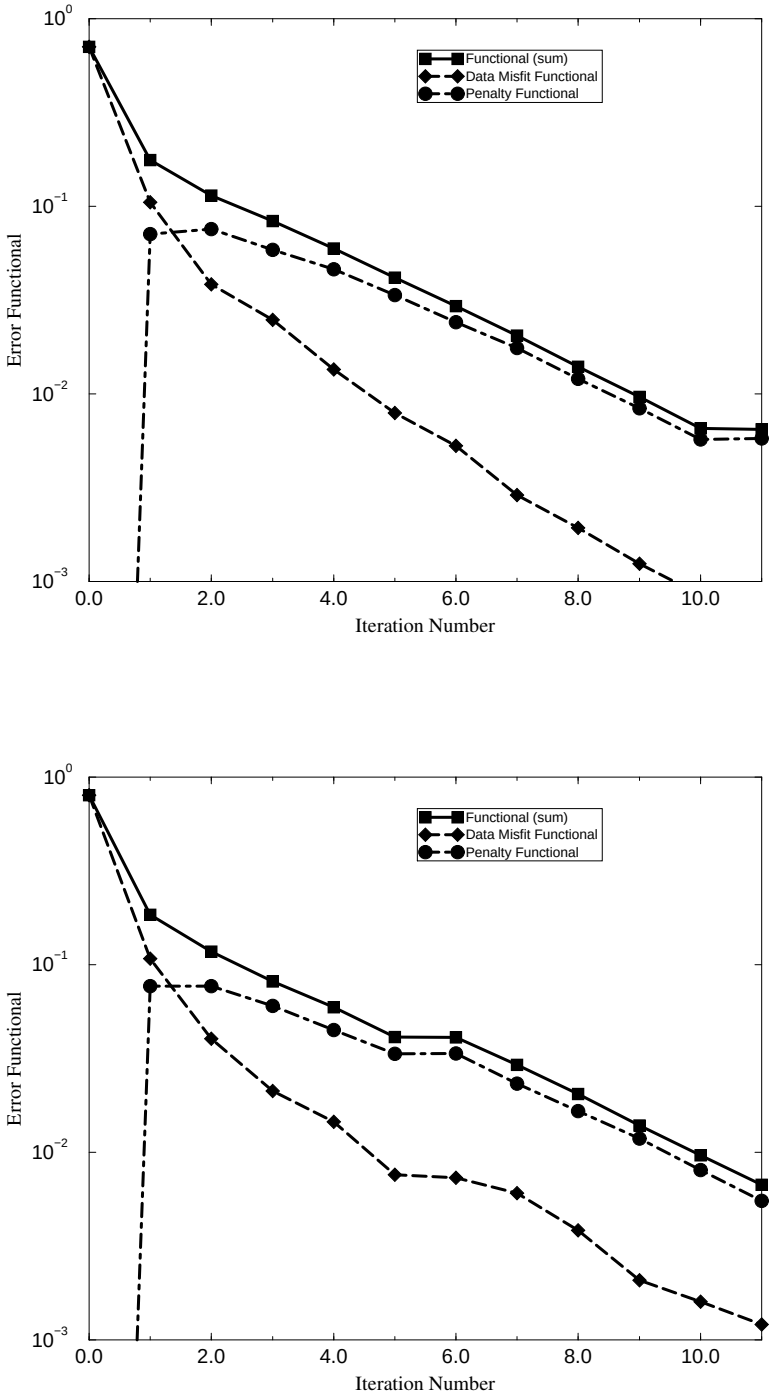


Figure 5.5: Error functional as a function of iteration number. For both P_1 (top graph) and P_3 (bottom graph) angular approximations the error functional is reduced most dramatically during early stages of the inversion. The penalty contribution to the error functional takes greatest effect between iterations 2 and 6. During these iterations deeply embedded material structure emerges.

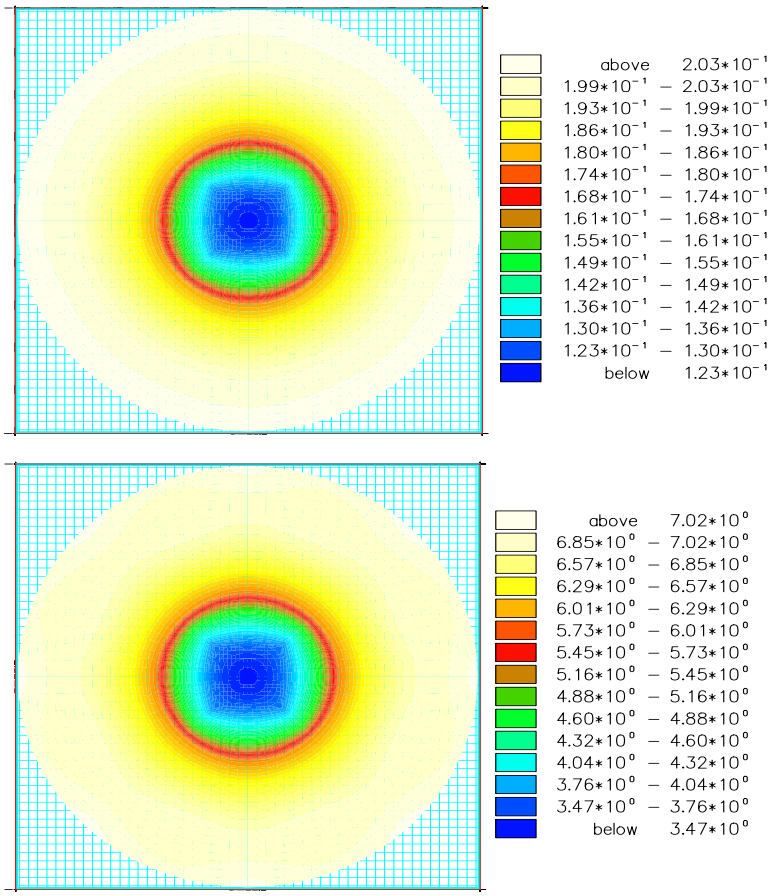


Figure 5.6: Absorption (top image) and scattering (bottom image) reconstruction results for a circle in circle type problem using a P_1 angular expansion and a Levenberg-Marquardt optimisation method. The problem geometry is the same as the problem geometry used to reconstruct figure 5.4 but a refined finite element mesh is used. Convergence was reached after 15 Levenberg-Marquardt iterations.

coefficients are also approximated fairly accurately. The P_1 and P_3 results approximate the magnitude of the absorption coefficient, μ_a, to extend from 0.2cm^{-1} to approximately 0.12cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 7.0cm^{-1} to approximately 3.4cm^{-1}. The magnitudes of both the scattering and absorption coefficients of the inhomogeneity are slightly overestimated, while the magnitudes of the background medium remain accurate.

Figure 5.5 illustrates the error functional as a function of iteration number using P_1 (top graph) and P_3 (bottom graph) angular approximations. Both the P_1 and P_3 angu-

lar approximations converged after 11 iterations to similarly small error functionals of $F = 0.00647$ and $F = 0.00672$ respectively. The convergence criterion is chosen so that once convergence is met, increasing the number of iterations does not change the inversion model significantly. For both P_1 and P_3 angular approximations the error functional is reduced most dramatically during early stages of the inversion (i.e. during the first four iterations). The penalty contribution to the error functional takes greatest effect between iterations 2 and 6. During these iterations deeply embedded material structure emerges.

Figure 5.6 compares the converged absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 5.2 and table 5.3. Calculations were performed using a P_1 angular expansion and a Levenberg-Marquardt optimisation technique. While the problem geometry and optimisation conditions remain largely the same as the conditions described in figure 5.4, the mesh used for the forward calculation and the inversion reconstruction was much more refined. The new finite element mesh had 4674 nodes and 6976 linear quadrilateral elements. The absorption and scattering penalty levels (γ_a and γ_{s0}) are also increased to an initial value of 1,000.

It is necessary to increase the regularisation penalty levels because refined meshes have more material property unknowns, and are therefore more sensitive to the effects of ill-posedness. Figure 5.6 illustrates that little difference is observed in both the absorption and scattering reconstruction results when a finite element mesh with more nodes is considered. The reconstruction converged after 15 iteration to an error functional of $F = 0.0056$. Four extra iterations were necessary because larger regularisation penalty levels are initially used.

Non-Linear Conjugate Gradient Method Versus the Levenberg-Marquardt Method

The performance of the Levenberg-Marquardt method is compared alongside the performance of the non-linear conjugate gradient method [72]. Medical optical tomography provides an inversion scheme that is generally regarded as highly non-linear and ill-posed. In other words the inversion scheme may produce non-unique solution's for a given inversion problem when insufficient data is supplied to the modelling domain. The non-linear

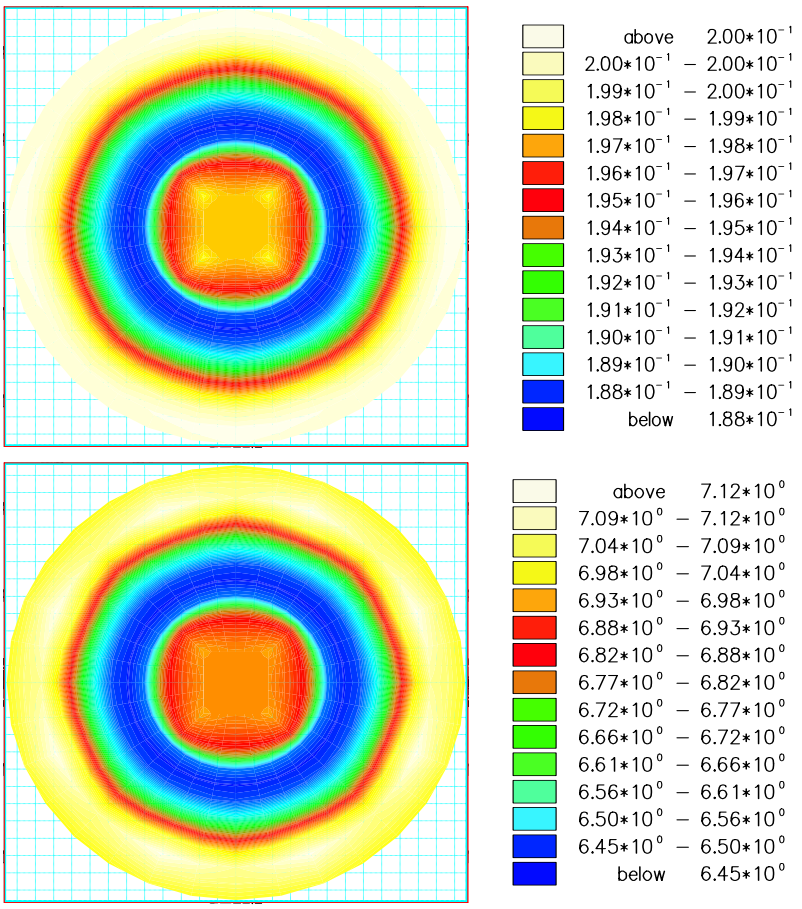


Figure 5.7: Absorption (top image) and scattering (bottom image) reconstruction results for a circle in circle problem using the non-linear conjugate gradient method. The inversion was performed for 100 non-linear conjugate gradient iterations using a P_1 angular expansion.

conjugate gradient method [72] is guided by the gradient of the error functional. Aside from being minimised at a solution's minimum the gradient of the error functional can also equal zero at maximum values and saddle points, thus providing a series of non-unique solutions. In contrast the Levenberg-Marquardt method is guided by both the formulation of the gradient vector and the approximate Hessian matrix, thus providing a more stable means of providing a solution.

Figure 5.7 illustrates the reconstructed material property distributions of the absorption coefficient (top image) and the scattering coefficient (bottom image) for the example outlined in figure 5.2 and table 5.3. The inversion was performed for 100 non-linear conju-

gate gradient iterations using a P_1 angular expansion. The approximate magnitude of the absorption coefficient, μ_a, is virtually uniform, extending from 0.2cm^{-1} to approximately 0.189cm^{-1}, while the magnitude of the scattering coefficient extends from 7.0cm^{-1} to approximately 6.45cm^{-1}. The absolute values of both the scattering and absorption coefficients of the inhomogeneity are over-estimated, whilst the approximate location and size of the inhomogeneity is also uncertain. On inspection figure 5.7 suggests that the inhomogeneity is a ring structure placed in the middle of the domain rather than a circle placed at the centre of the domain.

The relative failure of the non-linear conjugate gradient method when compared to the Levenberg-Marquardt method can be explained according to the way in which the non-linear conjugate gradient method treats regularisation terms. The non-linear conjugate gradient method is guided primarily by the gradient of the error functional. Material properties are therefore mainly adjusted around the sources and detectors, since it is here that the gradient sensitivities are largest. During the first non-linear iteration the gradient contains no regularisation contribution since the initial model is homogeneous. Therefore the model after the first non-linear conjugate gradient iteration is not influenced by regularisation and any 'damage' caused by a step in a structurally unregularised direction remains throughout the inversion. This is illustrated as the number of iterations is increased.

In contrast the Levenberg-Marquardt method treats regularisation terms implicitly. The implicit nature of the Levenberg-Marquardt method therefore enables structurally regularised inversion so that deeply embedded objects may be resolved even during the first Levenberg-Marquardt iteration. Regularisation parameters prevent material properties being adjusted primarily around the sources and detectors, while large regularisation parameters result in uniform material property updates. The Levenberg-Marquardt method converged after 11 iterations, reducing the error functional by two orders of magnitude from $F = 0.708$ to $F = 0.00647$. The non-linear conjugate gradient ran for 100 iterations and reduced the error functional to $F = 0.319$. The presented example illustrates, that inversion methods that treat regularisation implicitly, are better suited to finding the class of model requested by the structural constraints.

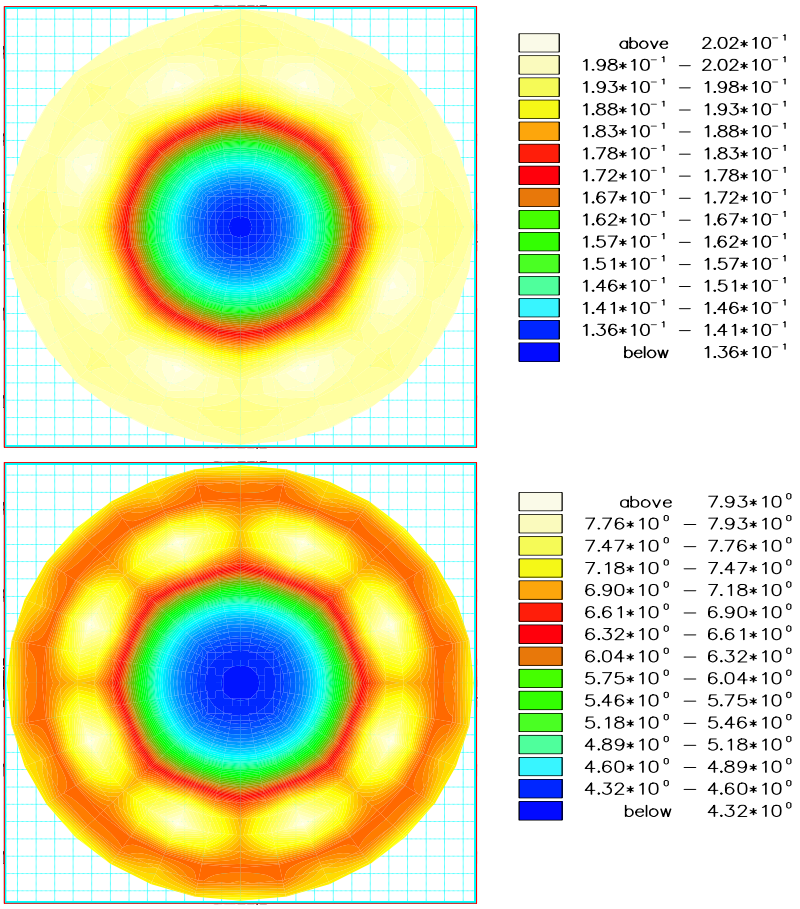


Figure 5.8: Absorption (top image) and scattering (bottom image) reconstruction results for a circle in circle type problem using a P_1 angular expansion. The inversion was performed for 14 Levenberg-Marquardt iterations without regularisation constraints.

Regularised and Unregularised Inversion Using the Levenberg-Marquardt Method

The regularisation functional, F_r, is used to impose structural constraints on a given model. The structural constraints are contained in the model covariance matrix, $\mathbf{K}$ (see equation 5.32). Performing an unregularised inversion is equivalent to setting the model covariance matrix, $\mathbf{K}$, equal to zero.

Figure 5.8 illustrates the Levenberg-Marquardt reconstruction of the absorption (top image) and scattering coefficients (bottom image) illustrated in figure 5.2 and table 5.3 without the application of structural constraints. The reconstruction was performed using a P_1 angular expansion. The relative magnitude of the scattering and absorption coefficients

are not approximated as accurately as the absorption and scattering coefficients resolved in figure 5.4. The approximate magnitude of the absorption coefficient, μ_a, extends from 0.2cm^{-1} to approximately 0.136cm^{-1}, while the magnitude of the scattering coefficient extends from 7.93cm^{-1} to approximately 4.32cm^{-1}. While the absorption reconstruction provides a reasonable inversion result the scattering coefficient is largely overestimated in the vicinity of the sources and detectors, thus providing the modelling domain with eight image artefacts. The unregularised inversion does not find a good model reconstruction because, like the non-linear conjugate gradient method, the inversion is largely guided by the gradient of the error functional. The gradient is largest around the sources and detectors, therefore the largest adjustments made to the material properties are made at the surface where the sources and detectors are located. This initially decreases the data misfit, but prohibits model updates at depth during later iterations.

During the first 14 Levenberg-Marquardt iterations, the error functional is reduced by approximately two orders of magnitude, from $F = 0.708$ to $F = 0.00124$. Although the error functional is markedly reduced the final reconstruction does not resemble the true model. As well as partially reconstructing the inhomogeneity placed deep inside the medium, structure in the inversion model is also concentrated near the surface of the domain where the sensitivity of the error functional to changes in material properties is largest.

Applying model covariance constraints turns an otherwise ill-posed problem into an over determined system of equations that can be easily solved using a Levenberg-Marquardt optimisation scheme. The nature and size of the model covariance penalty parameters determines the appearance of the inversion model. The size of the regularisation penalty parameters γ_a and γ_{s0} (equation 5.56) are chosen at each iteration of the solution of the inverse problem. Generally, it is assumed that the data misfit improves as the magnitudes of γ_a and γ_{s0} are decreased and model covariance constraints become relatively less important in the error functional. However, using values that are too small can cause detrimental effects to the model update, manifested in the over fitting of data in regions where sensitivities are highest (i.e. in regions where source-detector separations are small).

Figure 5.9 illustrates a Levenberg-Marquardt reconstruction of absorption coefficients

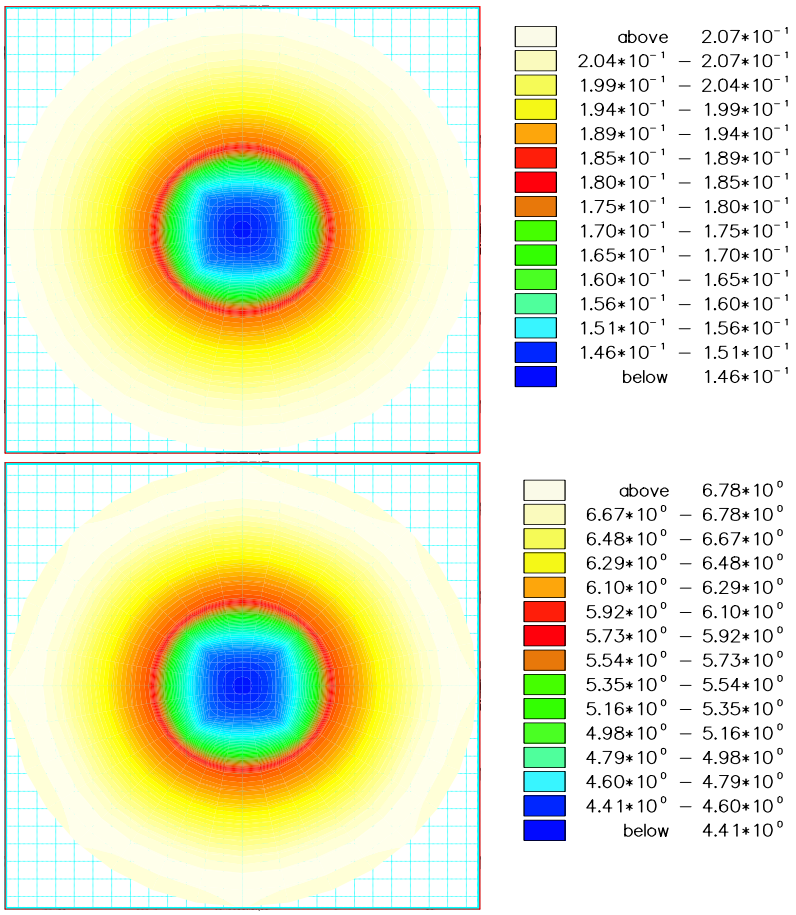


Figure 5.9: Absorption (top image) and scattering (bottom image) reconstruction results for a circle in circle type problem using a P_1 angular approximation. The inversion was performed for 14 Levenberg-Marquardt iterations using constant regularisation constraints of $\gamma_a = \gamma_{s0} = 100$.

(top image) and scattering coefficients (bottom image) for a system with a fixed regularisation penalty level of $\gamma_a = \gamma_{s0} = 100$. The reconstruction was performed using a P1 angular expansion for the geometry described in figure 5.2 and table 5.3. Figure 5.9 provides visual evidence that the algorithm finds the correct location and approximate size of the inhomogeneity for both absorption and scattering reconstructions. The relative magnitudes of the absorption and scattering coefficients are also approximated fairly accurately. The approximate magnitude of the absorption coefficient, μ_a, extends from 0.2cm^{-1} to approximately 0.146cm^{-1}. The magnitude of the scattering coefficient extends from 6.7cm^{-1} to approximately 4.41cm^{-1}.

Models resulting from large regularisation penalties stay smooth. As the penalty decreases, successively more structure develops. Models with small regularisation penalty values result in a rough inversion model, which sometimes does not bare resemblance to the true model. Figure 5.10 compares the error functional as a function of iteration number for the first 100 iterations of the non-linear conjugate gradient method (top graph) and calculations performed using the non-linear conjugate gradient method (dashed line with triangular symbols), the unregularised Levenberg-Marquardt method (the dotted line with diamond symbols), the regularised Levenberg-Marquardt method with constant regularisation penalties (the dashed line with circular symbols) and the regularised Levenberg-Marquardt method with a penalty relaxation equal to 1.5 during each iteration (the solid line with square symbols). The bottom graph illustrates decreases in the error functional during the first 11 iterations.

Figure 5.10 show that the non-linear conjugate gradient method decreases the error functional very slowly when compared to the Levenberg-Marquardt method illustrated in figure 5.5. After 100 non-linear conjugate gradient iterations the error functional is reduced to $F = 0.319$. This is a great deal larger than the error functionals obtained during Levenberg-Marquardt iterations. The bottom graph, illustrated in figure 5.10 suggests that in terms of data misfit or value of the error functional an unregularised model might be regarded as a 'good' inversion scheme, despite absorption and scattering image reconstructions suggesting the contrary. According to figure 5.10, the unregularised model provides the steepest decline in error functional and therefore the strongest convergence despite providing a scattering reconstruction showing eight image artefacts (figure 5.7) situated close to each of the sources and detectors. Figure 5.10 also illustrates that inversion performed with large or fixed penalty parameters generally result in large error functionals.

Although the reconstruction scheme illustrated in figure 5.9 provides visually pleasing results, the motivation for using relaxation of penalty parameters is twofold. Firstly, it is anticipated that by using large penalty parameters during early iterations, large scale structures are recovered. During later iterations, when the penalty parameters become smaller, small scale features are created. It is necessary to reconstruct small detail when considering complicated model geometries such as MRI images of the human brain. Sec-

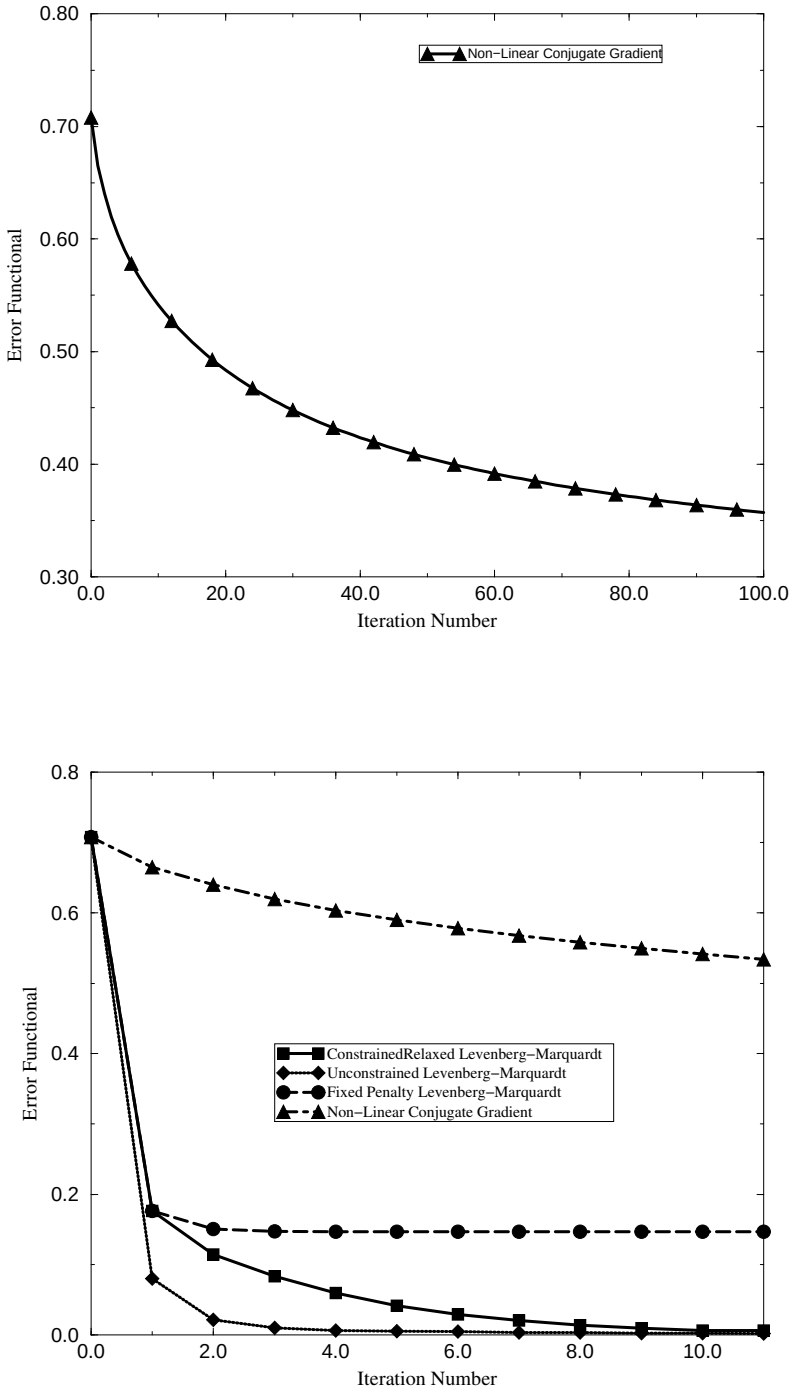


Figure 5.10: Error functional as a function of iteration number during non-linear conjugate gradient iterations (top graph) and calculations performed using the non-linear conjugate gradient method, the unregularised Levenberg-Marquardt method, the regularised Levenberg-Marquardt method and the regularised Levenberg-Marquardt method with a penalty relaxation equal to 1.5 during each iteration. The bottom graph illustrates the decrease in the error functional during the first 11 iterations.

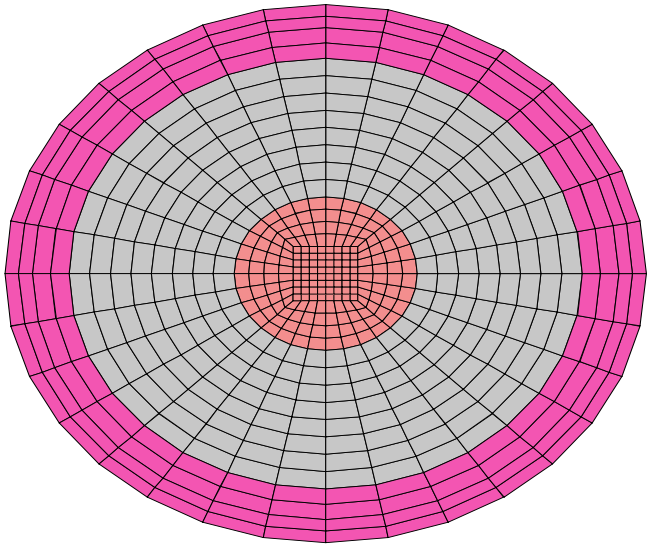


Figure 5.11: A two dimensional layered finite element circular mesh of diameter 7cm. The outer layer has a thickness of 0.7cm, while an inhomogeneity of diameter 2cm is located at the centre of the domain. The mesh used for the forward model and inversion model reconstruction has 2210 nodes and 3296 bi-linear quadrilateral elements.

ondly, it is difficult to specify optimum penalty levels without knowledge of error levels in the data and the amount of structure contained in the model. Thus it is anticipated that by relaxing the penalty levels after each iteration from a large initial value, an optimum penalty level is encountered somewhere along the way that will terminate the inversion process.

5.6.2 Experiment 2

Figure 5.11 illustrates a symmetric, layered circle. The circular test object is 7cm in diameter and contains a 0.7cm thick outer layer, an inner layer 1.8cm thick and an embedded inhomogeneity located at the centre of the domain. The inhomogeneity has a 2cm diameter. The optical parameters of each layer and the inhomogeneity are listed in table 5.4. Four sources and eight detectors are placed around the circumference of the medium to obtain the forward model and reconstruction model results. The sources are placed at co-ordinates (0.0, 3.5), (3.5, 0), (0.0, -3.5) and (-3.5, 0.0). Four of the detectors are

Regions	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Outer Layer (Background Medium)	0.2	7.0
Second Layer	0.05	2.0
Embedded Inhomogeneity	0.1	5.0

Table 5.4: Optical absorption and scattering coefficients for a layered circle 7cm in diameter containing a 0.7cm thick outer layer, an inner layer 1.8cm thick, and an embedded inhomogeneity located at the centre of the domain. The inhomogeneity has a 2cm diameter.

placed at the same co-ordinates, while the other four detectors are placed at co-ordinates (-2.5, 2.5),(2.5, 2.5), (2.5, -2.5) and (-2.5, -2.5). A P_1 angular approximation is used to obtain the steady state forward and inversion model calculations alongside a Levenberg-Marquardt optimisation method. The finite element mesh used for both the forward model and inverse reconstruction results has 2210 nodes and 3296 bi-linear quadrilateral elements. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium was chosen with optical properties equal to those of the background medium. In other words $\mu_a = 0.2$cm and $\mu_s^{'} = 7.0$cm.

Figure 5.12 compares the converged absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 5.11 and table 5.4. Since the optical parameters meet the criterion necessary to assume a highly scattering medium, a P_1 angular approximation provides an adequate description for modelling radiation transport. Figure 5.12 provides visual evidence that the inversion algorithm finds the correct location and approximate size of the inhomogeneity for both absorption and scattering reconstructions. The algorithm also illustrates the approximate layering thicknesses and provides good optical property magnitudes for each region.

The P_1 results approximate the magnitude of the absorption coefficient, μ_a, to extend from 0.162cm^{-1} to approximately 0.0648cm^{-1}, while the magnitude of the scattering

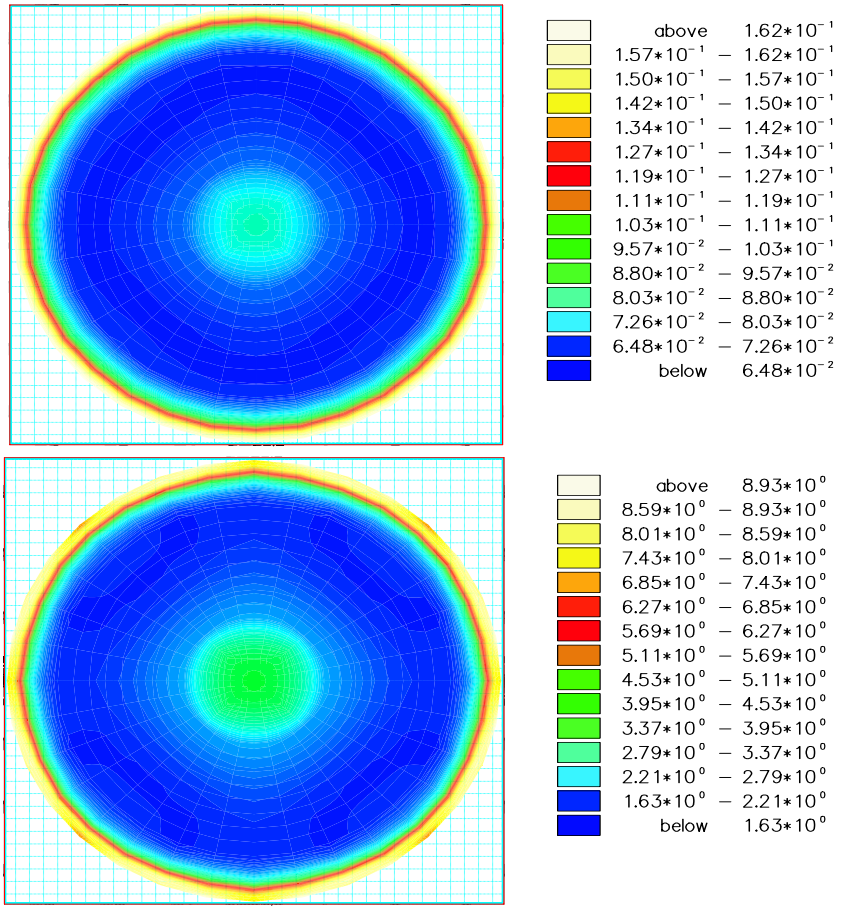


Figure 5.12: Absorption (top image) and scattering (bottom image) reconstruction results for a layered circular type problem using a P_1 angular approximation. The inversion converged after 20 Levenberg-Marquardt iterations.

Regions	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Outer Layer (Background Medium)	0.096 - 0.16	4.53 - 8.93
Second Layer	0.065 - 0.073	1.63 - 2.21
Embedded Inhomogeneity	0.083 - 0.10	3.95 - 5.11

Table 5.5: Absorption and scattering reconstruction optical properties for a layered circle 7cm in diameter containing a 0.7cm thick outer layer, an inner layer 1.8cm thick and an embedded inhomogeneity located at the centre of the domain. The inhomogeneity has a 2cm diameter.

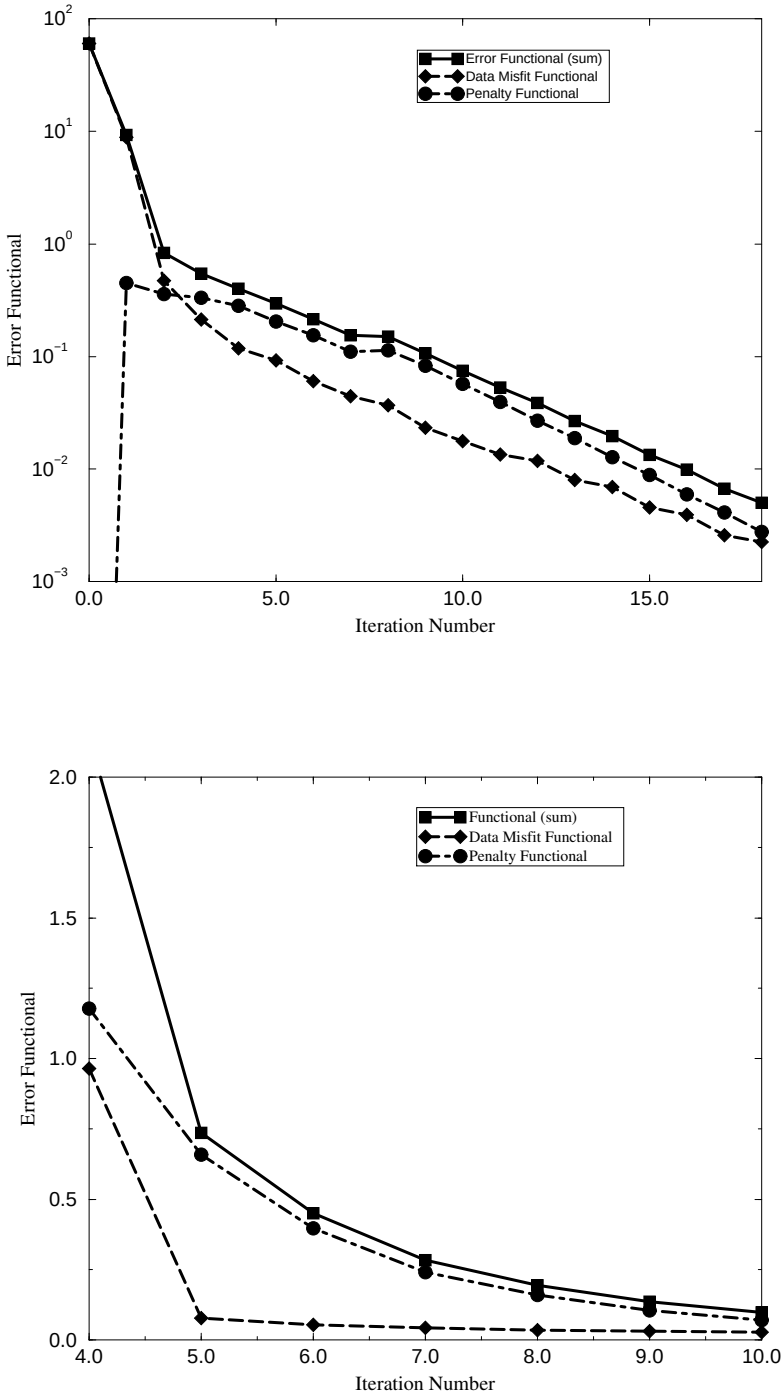


Figure 5.13: Error functional as a function of iteration number. The top graph illustrates the reduction of the error functional during the first 20 Levenberg-Marquardt iterations. The bottom graph extends from iteration number 4 to iteration 10. The penalty contribution contained in the error functional takes greatest effect during this period.

coefficient extends from approximately 8.93cm^{-1} to approximately 1.63cm^{-1}. Table 5.5 describes the range of absorption and scattering coefficients for the outer 0.7cm layer, the inside 1.8cm layer and the embedded inhomogeneity described in figure 5.11. The average optical properties of each region approximate that of the absolute optical property magnitudes described in table 5.4. The magnitudes of both the scattering and absorption coefficients of the inhomogeneity are most accurately approximated, while the magnitude of the scattering coefficients for the layered regions fall within the range described in table 5.4. The absorption coefficient magnitudes are underestimated in both layered regions.

Figure 5.13 illustrates the error functional as a function of the iteration number for the P_1 angular expansion described above. The top graph illustrates the model converging after 20 iterations to an error functional of $F = 0.00623$. The convergence criterion is chosen so that once convergence is met, increasing the number of iterations does not change the inversion model significantly. Figure 5.13 indicates that the error functional is reduced most dramatically during early stages of the inversion (i.e. during the first four iterations). The bottom graph illustrated in figure 5.13 illustrates that the penalty contribution to the error functional takes greatest effect between iterations 4 and 10. During these iterations fine, detailed structure emerges from deep within the model.

5.6.3 Experiment 3

Figure 5.14 illustrates a square domain of size 5cm $\times$ 5cm. The domain contains three embedded inhomogeneities, each of size 1cm $\times$ 1cm . The inhomogeneities are placed 1cm deep inside the medium. The absorption and scattering material properties of both the background medium and the embedded inhomogeneities are described in table 5.6. Two of the embedded inhomogeneities have the same absorption material properties and two of the embedded inhomogeneities have the same scattering material properties. Sixteen equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward and reconstruction model results. A P_1 angular approximation is used to obtain the steady state forward and inversion model calculations along with a Levenberg-Marquardt optimisation method. The same sized mesh is used to generate both the synthetic data and the inversion reconstruction. The mesh has 882

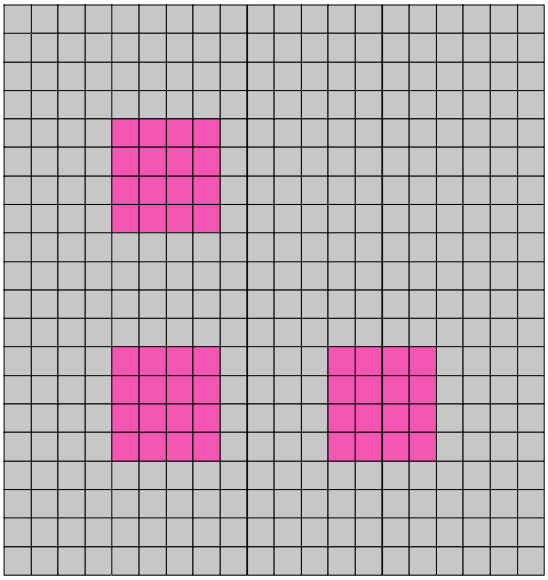


Figure 5.14: A two dimensional finite element square mesh for a phantom of size 5cm $\times$ 5cm. Three inhomogeneities are located 1cm deep inside the medium. The mesh used for the forward model and inversion model reconstruction has 882 nodes and 1280 bi-linear quadrilateral elements.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.07	10.0
Top Left Inhomogeneity	0.07	15.0
Bottom Left Inhomogeneity	0.14	10.0
Bottom Right Inhomogeneity	0.14	15.0

Table 5.6: Optical absorption and scattering coefficients for a square domain of size 5cm $\times$ 5cm. Three inhomogeneities are contained inside the domain. The scattering and absorption material properties of each of the inhomogeneities are described.

nodes and 1280 bi-linear quadrilateral elements. The absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100,000 and are annealed downwards by a factor of 1.5 after each iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium is chosen with optical properties that equal the background. In other words $\mu_a = 0.07$cm and $\mu_s^{'} = 10.0$cm.

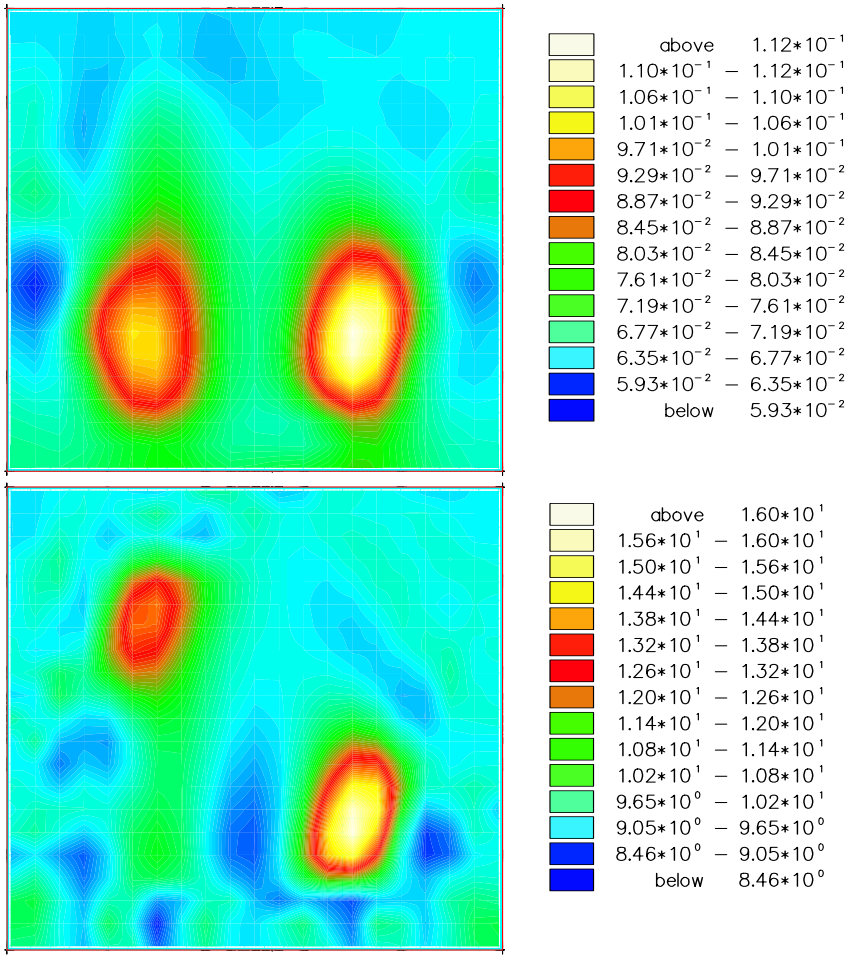


Figure 5.15: Absorption (top image) and scattering (bottom image) coefficient reconstruction results for a square domain problem containing three embedded inhomogeneities. Two of the embedded inhomogeneities have the same absorption material properties and two of the embedded inhomogeneities have the same scattering material properties. The inversion was performed using a P_1 angular approximation, a Levenberg-Marquardt optimisation scheme and was run for 43 iterations.

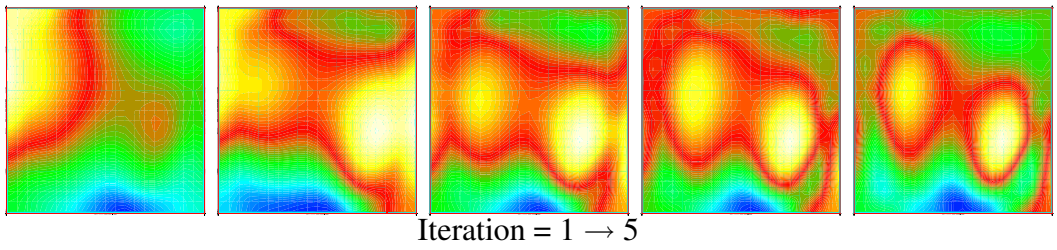
Figure 5.15 compares the converged absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 5.14 and table 5.6. Since the optical parameters meet the criterion necessary to assume a highly scattering medium, a P_1 angular approximation provides an adequate description for modelling radiation transport.

Figure 5.15 provides visual evidence that the inversion algorithm finds the correct loca-

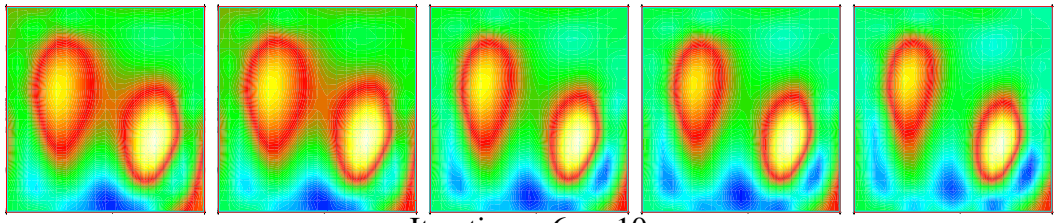
tion, size and approximate shape of the three inhomogeneities for both absorption and scattering reconstructions. The approximate magnitudes of the absorption coefficients and scattering coefficients of each of the three inhomogeneities are reconstructed relatively accurately. The P_1 results approximate the magnitude of the absorption coefficient, μ_a, to extend from 0.059cm^{-1} to approximately 0.112cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 16.0cm^{-1} to approximately 8.46cm^{-1}. The algorithm manages to distinguish between absorption and scattering features with relatively little cross-talk. The magnitude of the background absorption and scattering coefficients also remain accurate.

Figure 5.16 illustrates scattering (top set of images) and absorption (bottom set of images) reconstructions during the first 10 iterations of the Levenberg-Marquardt optimisation scheme. Starting from a homogeneous model, the basic material property structure emerges as the number of iterations is increased from 1 to 10. It is clear after 10 iterations that three inhomogeneities exist, two of which have the same scattering material properties and two of which have the same absorption material properties. After 10 iterations the magnitude of the absorption coefficient, μ_a, ranges from 0.0670cm^{-1} to approximately 0.0947cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 9.22cm^{-1} to approximately 12.0cm^{-1}. In contrast, after just 1 iteration the magnitude of both the absorption and scattering material properties remain fairly homogeneous. The magnitude of the absorption coefficient, μ_a, extends from 0.0747cm^{-1} to approximately 0.0809cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 10.0cm^{-1} to approximately 10.60cm^{-1}. The respective magnitudes of the scattering and absorption material properties remain homogeneous during early iterations because the magnitude of the regularisation penalty levels (γ_a and γ_{s0}) are initially very large.

At first glance, the results described appear contradictory to the results described by Arridge and Lionheart [37]. Arridge and Lionheart have shown that if the (steady state) diffusion equation is assumed as the model of light propagation in optical tomography, the simultaneous unique recovery of diffusion and absorption coefficients cannot be achieved [37]. In other words, there is no unique solution to the inverse problem. However, the model described utilises a P_1 angular approximation to describe diffusive radiation trans-

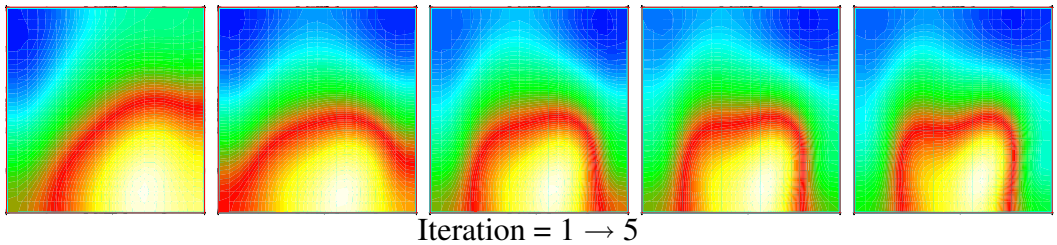


Iteration = 1 → 5

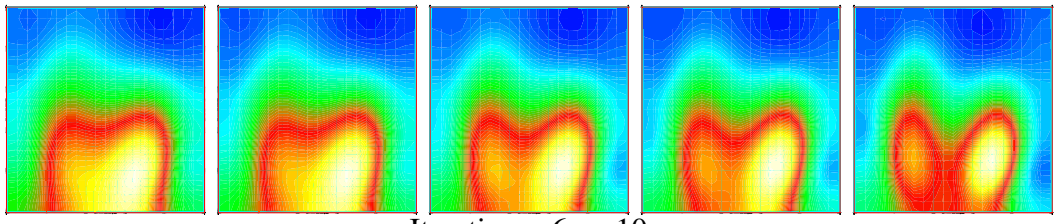


Iteration = 6 → 10

Scattering Coefficient Images



Iteration = 1 → 5



Iteration = 6 → 10

Absorption Coefficient Images

Figure 5.16: Scattering (top set of images) and absorption (bottom set of images) reconstruction results during the first 10 iterations of a square domain problem containing three inhomogeneities. The inversion was performed using a P_1 angular approximation and a Levenberg-Marquardt optimisation scheme. Notice the basic absorption and scattering material property structure emerging as the number of iterations are increased (left to right) from 1 to 10. Detailed material structure and accurate material property magnitudes are resolved during later iterations.

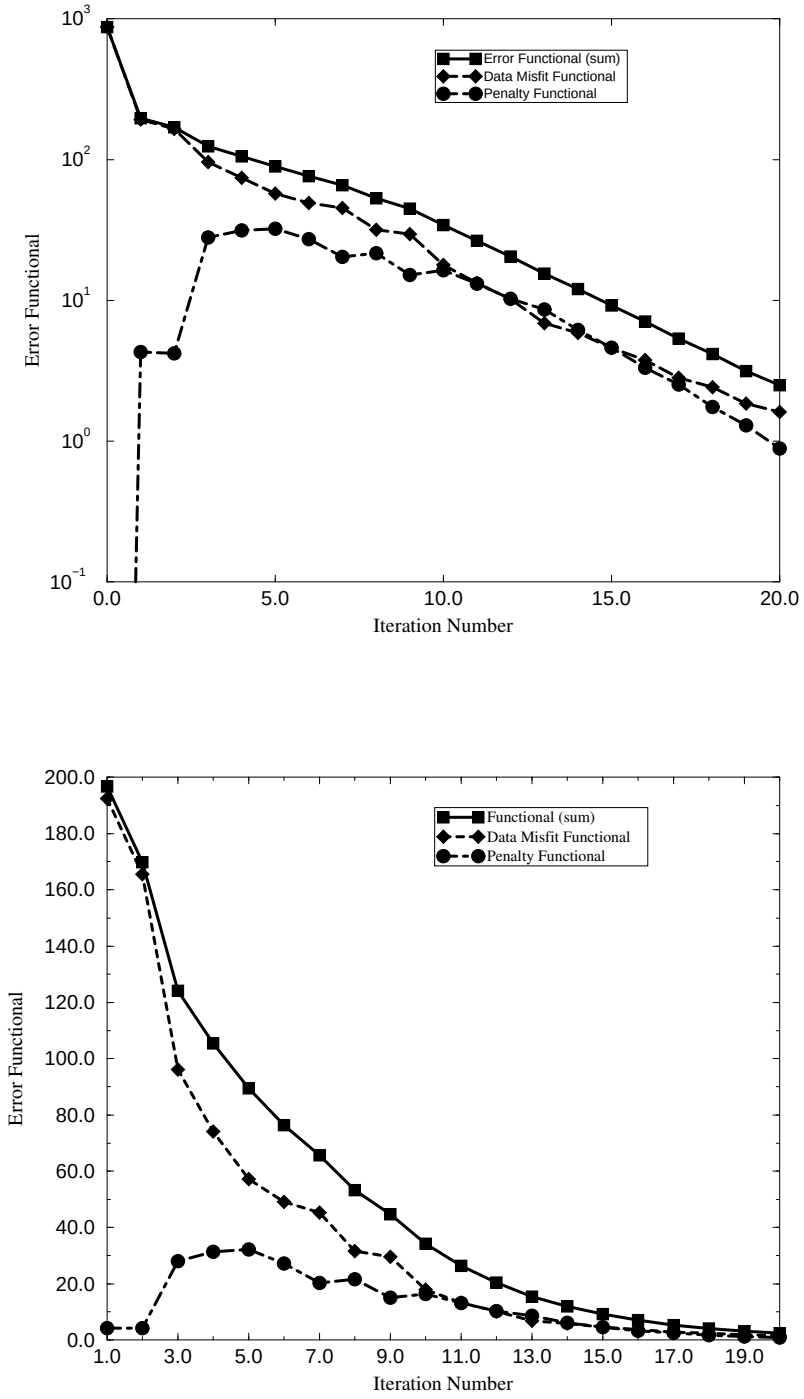


Figure 5.17: Error functional as a function of iteration number. The top graph illustrates the reduction of the error functional during the first 20 Levenberg-Marquardt iterations. The bottom graph also extends from iteration number 1 to iteration 20. The gradual decline of the error functional after the first Levenberg-Marquardt iteration is because γ_a and γ_{s0} are initially set to large magnitudes before being slowly annealed downwards.

port. The P_1 angular approximation is not entirely equivalent to the diffusion approximation. The diffusion approximation represents a hyperbolic system whereas a P_1 angular expansion represents a parabolic system. It was suggested by Matcher et al [170] that uniqueness may be restored if the model is taken to be a P_1 approximation. The above result may therefore provide additional evidence to support this claim.

Figure 5.17 illustrates error functional as a function of iteration number for the first 20 Levenberg-Marquardt iterations of the P_1 angular expansion. The model converged after 43 iterations from an error functional of $F = 873.855$ to an error functional of $F = 0.162$. Like figures, 5.5 and 5.13, figure 5.17 indicates that the error functional is reduced most dramatically during early stages of the inversion. However, the decline in the error functional is not as sharp as described in figures 5.5 and 5.13. This is because a high initial regularisation penalty parameter of 100,000 is selected for both γ_a and γ_{s0}. The high regularisation penalty results in an initially homogeneous model and provides a few fairly homogeneous material property updates. The regularisation penalty is annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The small relative reduction in γ_a and γ_{s0} provides a gradual decrease in the error functional. The inversion takes a greater number of iterations to converge because the regularisation is essentially taking its time and also because the functional is being reduced from a high initial value.

5.6.4 Experiment 4

The final simulation geometry, described in figure 5.18, is a square domain containing an embedded transparent, void-like ring. The ring has a 1.0cm thickness and is placed 1.0cm deep into the domain. The absorption and scattering material properties of both the background medium and the transparent void are described in table 5.7. Sixteen equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward and reconstruction model results. P_1, P_3, and P_5 angular approximations are used to obtain steady state forward and inversion calculations using a Levenberg-Marquardt optimisation method. The same sized mesh is used to generate both the synthetic data and the inversion reconstruction. The mesh has 882 nodes and 1280 linear quadrilateral elements. Both the absorption and scattering regularisation

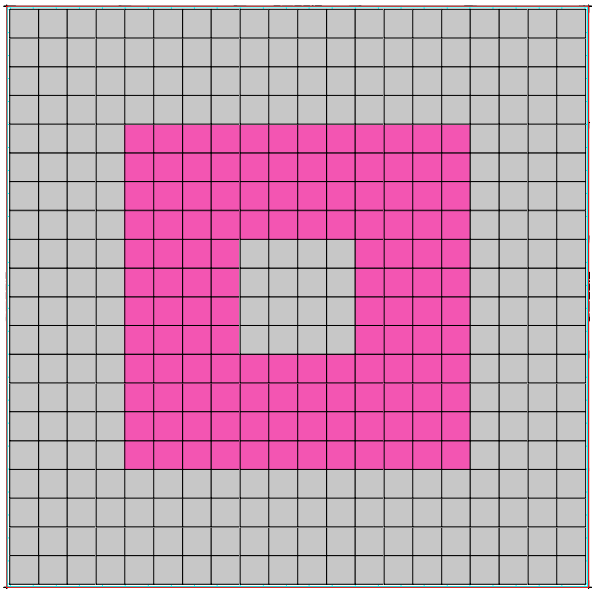


Figure 5.18: A two dimensional finite element square mesh for a phantom of size 5cm $\times$ 5cm. A void-like ring of thickness 1.0cm is located at a depth of 1.0cm into the domain. The mesh used for the forward model and inversion model reconstruction has 882 nodes and 1280 bi-linear quadrilateral elements.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.07	10.0
Ring Region	0.0	0.0

Table 5.7: Optical absorption and scattering coefficients for a square domain of size 5cm $\times$ 5cm. The square domain contains a transparent void region in the shape of a ring.

penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium was chosen with optical properties equal to those of the background medium. In other words $\mu_a = 0.07$cm and $\mu_s^{'} = 10.0$cm.

Figures 5.19 and 5.20 compare the absorption (top set of images) and scattering (bottom set of images) reconstruction results, for the problem geometry described in figure

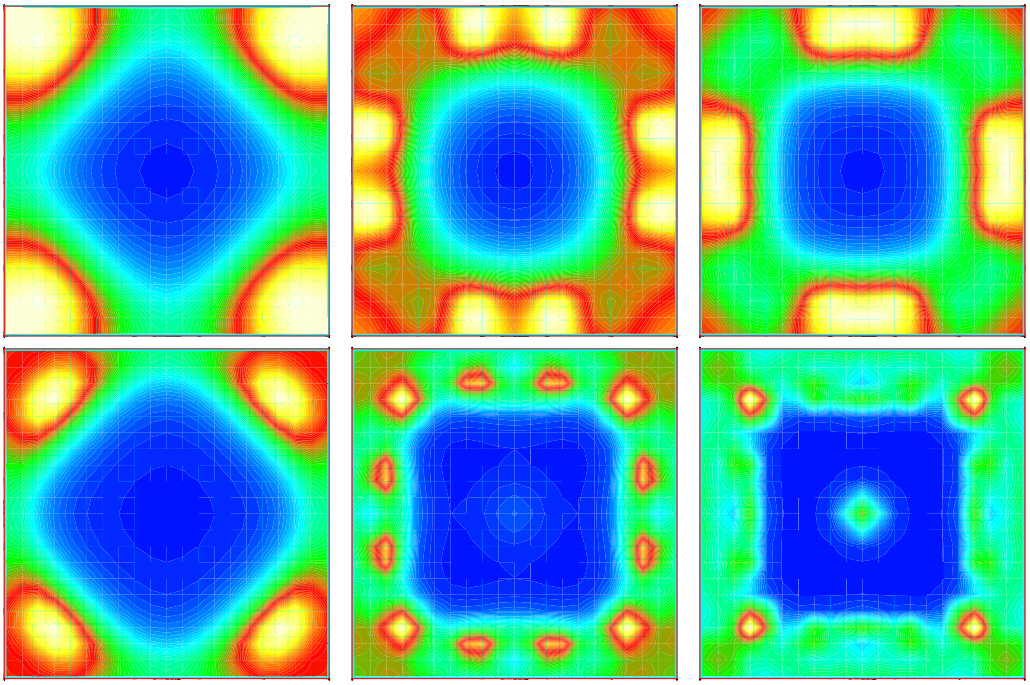


Figure 5.19: The absorption (top set of images) and scattering (bottom set of images) reconstruction results after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively, for the problem geometry described in figure 5.18 and table 5.7. A P_5 angular approximation was used to perform the inversion calculations together with a Levenberg-Marquardt optimisation scheme.

5.18 and table 5.7, after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively. Figure 5.19 illustrates a P_5 angular approximation result, while figure 5.20 illustrates a P_3 angular approximation result. Both figures 5.19 and 5.20 provide visual evidence that during the respective scattering reconstructions, the algorithm finds the correct location and approximate thickness of the transparent ring. However, both the P_3 and P_5 angular approximations fail to reconstruct the absorption coefficient material structure. Instead of reconstructing a ring-like region, a large square shaped void is reconstructed.

Figure 5.21 compares the absorption (top set of images) and scattering (bottom set of images) reconstruction results after (from left to right) 1 iteration, 10 iterations and 50 iterations for a P_1 angular expansion. The P_1 angular approximation fails to reconstruct both the scattering and absorption material structure. According to the P_1 images described in figure 5.21 the ring-like structure described in the original problem geometry

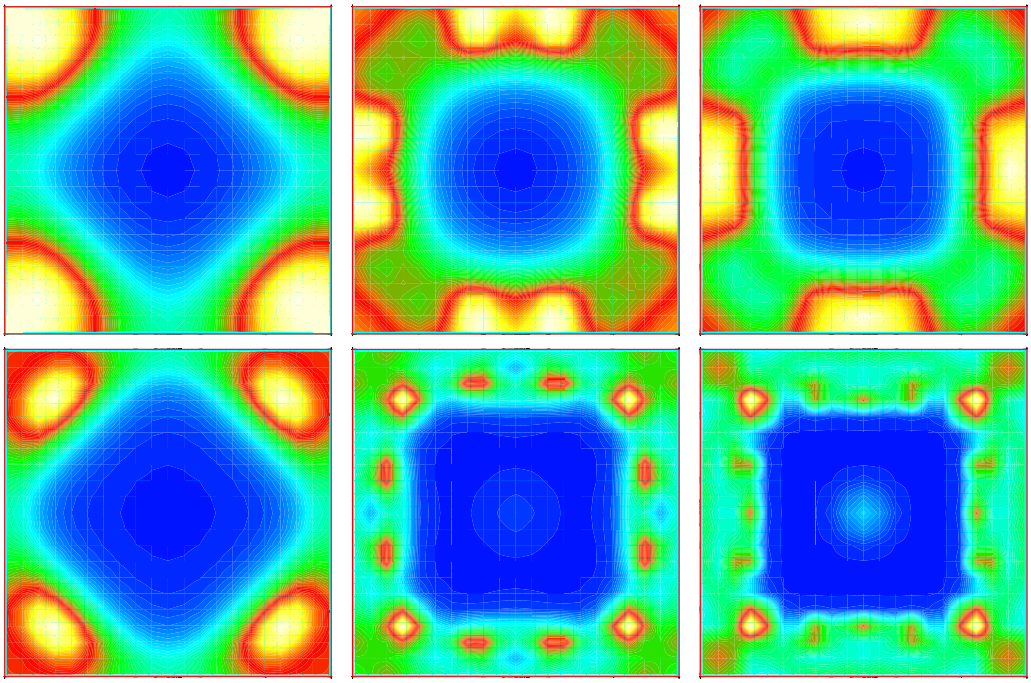


Figure 5.20: The absorption (top set of images) and scattering (bottom set of images) reconstruction results after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively, for the problem geometry described in figure 5.18 and table 5.7. A P_3 angular approximation was used to perform the inversion calculations together with a Levenberg-Marquardt optimisation scheme.

(figure 5.18) does not exist. Instead a large void is reconstructed, which only resembles a square structure for the scattering reconstruction. The results confirm that for steady state, void-like configurations, in which diffusion theory is not valid, full transport theory is necessary to reconstruct an objects material structure.

The relative magnitudes of the absorption and scattering coefficients are approximated fairly accurately although they are never absolutely resolved. The P_1 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.182cm^{-1} to approximately 0.019cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.41cm^{-1} to approximately 0.213cm^{-1}. The P_3 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.102cm^{-1} to approximately 0.013 cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 10.0cm^{-1} to approximately 0.278cm^{-1}. Finally, the P_5 angular ex-

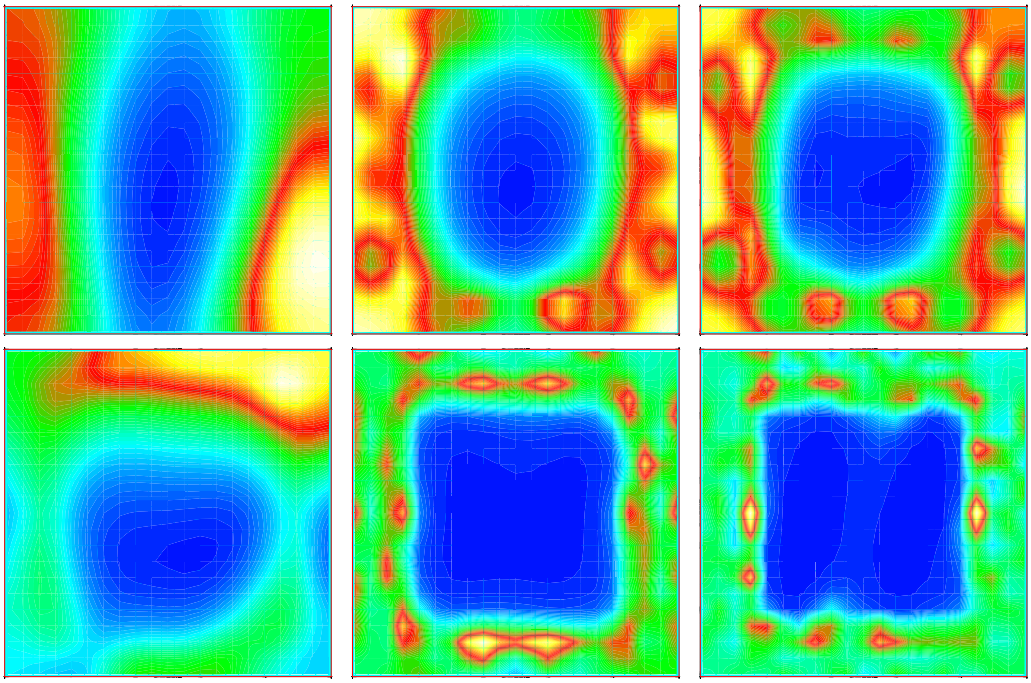


Figure 5.21: The absorption (top set of images) and scattering (bottom set of images) reconstruction results after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively, for the problem geometry described in figure 5.18 and table 5.7. A P_1 angular approximation was used to perform the inversion calculations together with a Levenberg-Marquardt optimisation scheme.

pansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.099cm^{-1} to approximately 0.012 cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 10.0cm^{-1} to approximately 0.30cm^{-1}. For each of the angular approximations (P_1, P_3 and P_5), the magnitudes of both the scattering and absorption coefficients of the inhomogeneous ring indicate that a void is present but are overestimated. The P_5 result also provides a significantly improved reconstruction of the internal scattering region when compared to the P_3 angular approximation results. During P_3 and P_5 angular approximations artefacts exist for both scattering and absorption reconstructions (e.g. the sixteen scattering coefficient peaks that have scattering coefficient extending up to 35.0cm^{-1}.

During the P_1 and P_3, and P_5 angular approximations the error functional is reduced by more than three orders of magnitude. The P_1 angular expansion reduces the error

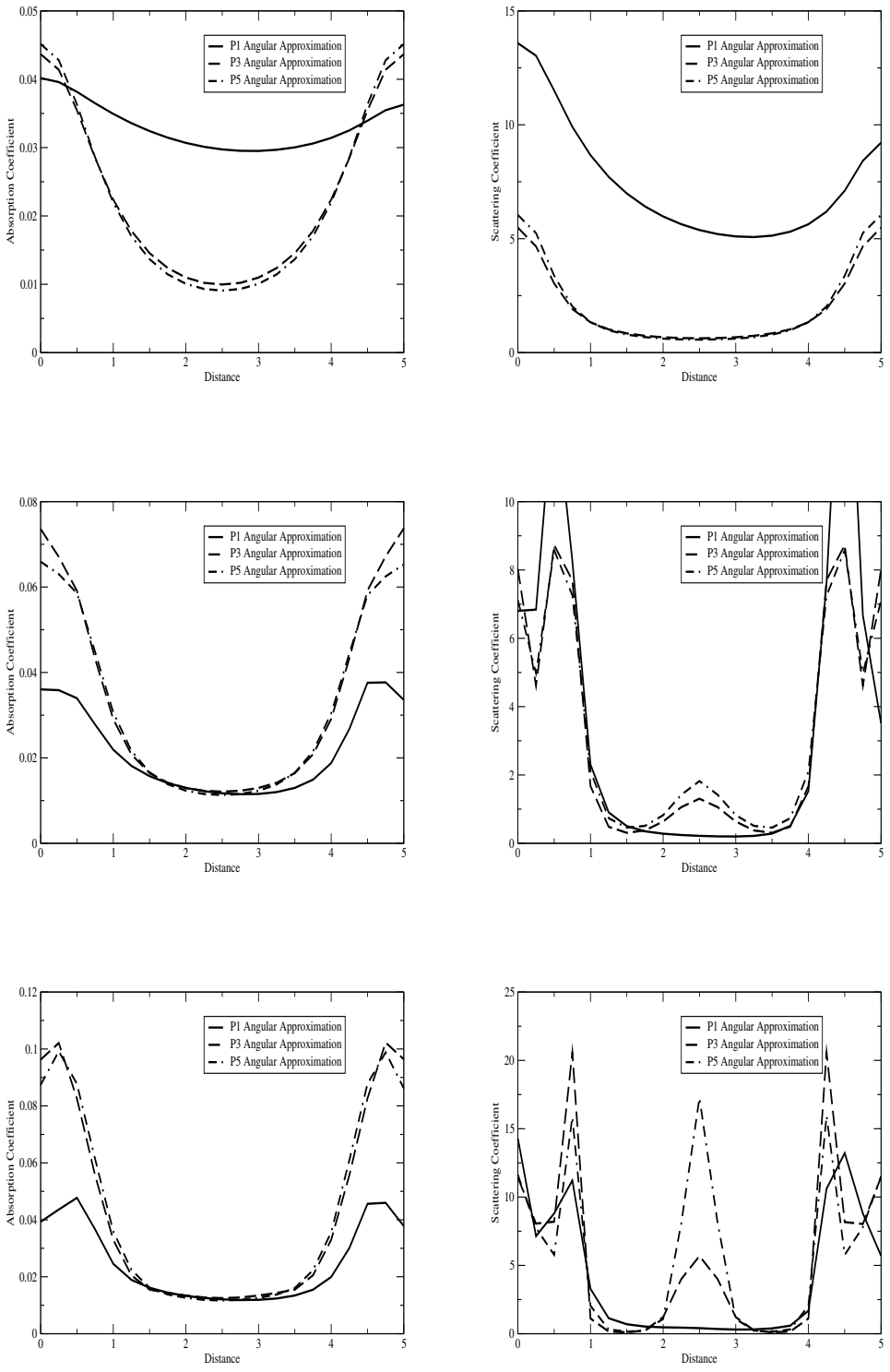


Figure 5.22: Absorption coefficient (left hand graphs) and scattering coefficient (right hand graphs) as a function of distance along the x-axis for the P_5, P_3 and P_1 angular approximations described in figures 5.19, 5.20 and 5.21 after 1 iteration (top graph), 10 iterations (middle graph) and 50 iterations (bottom graph) respectively. The dotted line represents the exact solution.

functional from $F = 37273$ to $F = 33.07$. The P_3 angular approximation reduces the error functional from $F = 33843$ to $F = 25.97$ and the P_5 angular approximation reduces the error functional from $F = 32390$ to $F = 23.16$.

Figure 5.22 depicts a graph of absorption coefficient (left hand side graphs) and scattering coefficient (right hand side graphs) as a function of distance along the x-axis for the P_5, P_3 and P_1 angular approximations described in figures 5.19, 5.20, and 5.21 respectively. A line is drawn through the middle of the y-axis, at co-ordinates (0.0, 2.5), (5.0, 2.5). The top graphs illustrates absorption coefficient and scattering coefficient as a function of distance after the first iteration. The middle graph illustrates absorption coefficient and scattering coefficient as a function of distance after the tenth iteration and the bottom graph illustrates absorption coefficient and scattering coefficient as a function of distance after the fiftieth iteration. The dotted line represents the exact solution.

The absorption coefficient reconstructions illustrated in figure 5.22 shows that the void is largely resolved during the first iteration. After 1 iteration the absorption coefficient extends from approximately $\mu_a = 0.044\text{cm}^{-1}$ to approximately $\mu_a = 0.01\text{cm}^{-1}$, while after 50 iterations the absorption coefficient extends from approximately $\mu_a = 0.1\text{cm}^{-1}$ to approximately $\mu_a = 0.01\text{cm}^{-1}$. Figure 5.22 illustrates that, even after 50 iterations, the ring is never reconstructed for absorption material properties for each of the angular approximations. Like the absorption coefficient reconstruction, during scattering coefficient reconstructions the void is largely established after the first iteration. However, further structure is reconstructed between iterations 1 and 10. After the tenth iteration the ring-like structure emerges for both P_3 and P_5 angular approximations and is illustrated in figure 5.22 as a peak that extends between a distance of $x = 2$cm and $x = 3$cm. The presence of the ring structure is confirmed in the spatial reconstruction illustrated in figures 5.19, and 5.20. After 50 iterations the ring is fully reconstructed and has the correct approximate thickness and scattering coefficient magnitude. Two of the sixteen image artefacts are also reconstructed during scattering reconstructions. The peaks have scattering coefficients extending to approximately $\mu_a = 21\text{cm}^{-1}$.

5.7 Conclusion

A main contribution of this chapter is the development of a differentiable discretisation of the Boltzmann transport equation. This discretisation has been designed specifically for inverse problems and accounts for both optically thick (diffusive) and transparent media. Another contribution is the overall inversion method that has been outlined above.

Four synthetic problems are used to demonstrate the simultaneous reconstruction of absorption and scattering material properties inside a host medium. While, the phantom problems are simple it is hoped that the inversion method will eventually be applied to real life medical imaging scenarios. The examples illustrate the need to use full transport theory when considering void regions. The examples also illustrate the need to consider regularisation matrices when dealing with ill-posed inversion problems. Furthermore, the importance of treating the regularisation terms implicitly has also been demonstrated. This was achieved by comparing the performance of the Levenberg-Marquardt method, with a non-linear conjugate gradient method. The latter method performed poorly because of the explicit treatment of regularisation. Relaxing the magnitude of the regularisation penalty levels from large to small values during the inversion resulted in the most accurate inversion results. The necessary efficiency for the computationally intensive problems described has been achieved with appropriate preconditioning and adaptive control of the diagonal dominance of the least squares problem.

Chapter 6

Finite Element Optical Imaging Method Using Transport Theory and Time Dependent Information

Contents

The following chapter focuses on the transport of photons during medical optical tomography. The transport of near infrared optical radiation is described by the one-speed, time-dependent Boltzmann transport equation. A Streamline Upwinded Petrov-Galerkin (SUPG) method is used to spatially discretise the Boltzmann transport equation and spherical harmonic basis functions are used to represent the angle of photon travel. SUPG methods are accurate and robust across all radiation transport regimes from optically thick to thin media and form a robust coupling between time and space dimensions. A space-time approach to the discretisation has been adopted using a Discontinuous Galerkin (DG) method for the time domain discretisation. The inversion capabilities are demonstrated using time dependent information.

6.1 Introduction

As well as providing a forward model that utilises full radiation transport theory, introducing time dependent information into the imaging domain can also improve forward model accuracy [50, 51]. The only way to guarantee an accurate calculation of the number of photons observed at a series of detectors is to resolve the forward model's transient signal [55, 56, 57]. However, in many areas of optical imaging, scattering dominant light transport often makes it difficult or even impossible to observe an accurate temporal signal. Factors that reduce signal quality include, noise, the small size of the time window in which photons are observed, and the relatively small number of photons observed at the detectors (e.g. when imaging the neonatal brain). The time-dependent signal is often

resolved using either a transform (.e.g. Laplace transform) or a truncated Fourier series expansion (in time) [57]. For time marching methods this often requires complex check pointing algorithms [171] or a large amount of computer memory. Furthermore, although Fourier series expansions have exponential convergence properties in the time domain, Fourier analysis requires the time domain to be periodic. Fourier series expansions also suffer heavily from Gibbs oscillations that result from abruptly changing time signals. The resulting linear equations tend to be poorly conditioned, thus reducing the effectiveness of the iterative linear solution of the forward and adjoint problems. Many groups choose to simplify the inversion algorithm by mapping the governing equations into the frequency domain [172]. This results in a discretisation of the decoupled equations for each frequency argument and is achieved using a Fourier series expansion.

A Streamline Upwinded Petrov-Galerkin (SUPG) technique [58, 59, 60, 61, 62] is used to discretise the Boltzmann transport equation in space and time. Angular discretisation is achieved using spherical harmonics. SUPG methods are accurate and robust across all radiation transport regimes from optically thick to thin media and form a robust coupling between the time and space dimensions. The SUPG approach already works for imaging with steady state signals (chapter 5) and has been specifically formulated for use with inverse problems. For inversion regimes, the SUPG method is designed to yield a set of discrete equations that can be differentiated with respect to an objects material properties. This allows gradients to be formed as part of the inversion procedure. The SUPG time stepping method has the ability to use large time steps while still capturing the essential features of the time dependent signals. The approach uses a Discontinuous Galerkin (DG) discretisation of the time domain. For each of the model geometries considered in this thesis, the photon velocity is set arbitrarily to 1 and a time step of $\Delta t = 5$ is used. The discretisation is second order accurate at the end of each time level, and third order at the beginning of each time level. Another desirable property of the chosen time discretisation is its ability to accurately resolve low frequencies while filtering out unresolved higher frequencies.

Scattering dominant light transport [63, 64, 65] is a limiting factor in the development of optical tomography. The progressive influence of scattering with distance makes the scattering problem complex and non-linear. Iterative inversion methods are therefore

considered to evaluate what optical properties would have created the original signal [3, 4, 26, 35, 40, 41, 42, 68, 69, 70]. A multi-dimensional inversion scheme has been developed to describe the transport of photons during near infrared optical tomography. The technique enables the reconstruction of optical property distribution in two and three dimensions for arbitrary model geometries and for arbitrary source and detector positions along the boundary of the domain. The inversion problem is formulated as a functional optimisation [41, 69, 74, 75, 76] that utilises an error functional containing a data misfit term and model covariance terms. The model covariance data measures both the smoothness of the model and any deviation from the starting model. The discretised form of the functional is minimised using a gradient based, second order Levenberg-Marquardt technique. Both the source problem for the angular flux and the least squares (inverse) problem are solved using an iterative (linear) preconditioned conjugate gradient solver in order to bypass the explicit computation of the Jacobian and Hessian matrices required for Levenberg-Marquardt iterations. Model covariance regularisation is initially exploited to provide uniform model updates before being intermittently annealed downwards so that structure emerges during later iterations. Iterations are performed until the predicted data agrees with the detector readings. The final distribution of optical properties is then displayed as an image.

The main contribution of this chapter is the development of a time dependent inversion scheme. A series of applications have been considered to demonstrate the enhanced imaging achievable using the time dependent form of the Boltzmann transport equation. The chapter is organised as follows. Section 6.2 describes the equations used to model time-dependent radiation transport alongside their finite element discretisation. Section 6.3 defines the error functional used in the inversion of non-linear ill-posed problems. The functional has terms representing data misfit and model covariance with a view of penalising both structure and deviation from the starting model. The optimisation of the error functional by means of a modified Levenberg-Marquardt method and a description of how calculations involving Jacobian and approximate Hessian matrices are approached is described in section 6.4, while the numerical applications described in section 6.5 demonstrate the feasibility of the proposed method. Finally, our findings are summarised in section 6.6.

6.2 Governing Equations

For optical tomography applications the time dependent, one-speed Boltzmann transport equation is used to describe the angular flux, $\psi(\mathbf{r}, \mathbf{\Omega}, t)$, that results from an external source of photons, $s(\mathbf{r}, \mathbf{\Omega}, t)$, incident on a host medium. The time dependent Boltzmann transport equation for mono-energetic photons is written:

$$\frac{1}{v}\frac{\partial \psi(\mathbf{r}, \mathbf{\Omega}, t)}{\partial t} + \mathbf{\Omega} \cdot \nabla \psi(\mathbf{r}, \mathbf{\Omega}, t) + \mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}, t) = s(\mathbf{r}, \mathbf{\Omega}, t), \tag{6.1}$$

where $\mathbf{r}$ is the position vector $\mathbf{r} = (x, y, z)$ of a particle or photon moving along direction, $\mathbf{\Omega} = (\theta, \omega)$, at time, t. The system has characteristic velocity, v. $\psi(\mathbf{r}, \mathbf{\Omega}, t)$ represents the particle angular flux and $s(\mathbf{r}, \mathbf{\Omega}, t)$ represents an external source term. The scattering/removal operator, $\mathcal{H}$ is defined:

$$\mathcal{H}\psi(\mathbf{r}, \mathbf{\Omega}, t) = (\mu_a(\mathbf{r}) + \mu_s(\mathbf{r}))\psi(\mathbf{r}, \mathbf{\Omega}, t) - \int_{\Omega} \mu_s(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega})\psi(\mathbf{r}, \mathbf{\Omega}', t)d\mathbf{\Omega}', \tag{6.2}$$

where $\mu_a(\mathbf{r})$ and $\mu_s(\mathbf{r})$ represent position dependent absorption and scattering coefficients respectively, and $\mu_s(\mathbf{r}, \mathbf{\Omega}' \rightarrow \mathbf{\Omega})$ represents the differential macroscopic cross-section. Boundary conditions must be defined for a given problem. Each of the models presented in this report have vacuum boundary conditions [34]:

$$\psi(\mathbf{r}_s, \mathbf{\Omega}, t) = 0, \quad when \quad \mathbf{n} \cdot \mathbf{\Omega} < 0, \tag{6.3}$$

in which, $\mathbf{n}$ represents the normal to the boundary. In other words, any photon that escapes through the bounding surface is not able to re-enter the medium through another part of the surface.

6.2.1 Finite Element Approximation of the Forward Model

In order to solve equation 6.1 numerically the source term is subtracted from both sides of equation 6.1 and the subsequent expression is discretised in angle and space. Angular discretisation is achieved using the Petrov-Galerkin method [58, 59, 119, 120]. The Petrov-Galerkin method is applied by multiplying the entire expression by a series of angularly dependent finite element weight functions, $\mathcal{M}_\rho(\mathbf{\Omega})$, and integrating the subsequent expression over the solution domain. Additionally a finite element approximation $\mathbf{\Psi}(\mathbf{r},t)$, replaces the continuous angular flux, $\psi(\mathbf{r},\mathbf{\Omega},t)$ and a finite element approximation, $\mathbf{S}(\mathbf{r},t)$, is also used to approximate the source function, $s(\mathbf{r},\mathbf{\Omega},t)$. The system of equations become:

$$\int_{\mathbf{\Omega}} M_\rho(\mathbf{\Omega}) \left[\frac{1}{v}\frac{\partial\psi(\mathbf{r},\mathbf{\Omega},t)}{\partial t} + \mathbf{\Omega}\cdot\nabla\psi(\mathbf{r},\mathbf{\Omega},t) + \mathcal{H}\psi(\mathbf{r},\mathbf{\Omega},t) - s(\mathbf{r},\mathbf{\Omega},t) \right] d\mathbf{\Omega} = 0. \tag{6.4}$$

The approximation $\mathbf{\Psi}(\mathbf{r},t)$ for the angular flux, $\psi(\mathbf{r},\mathbf{\Omega},t)$, is described using the finite element basis function, $M_\nu(\mathbf{\Omega})$:

$$\psi(\mathbf{r},\mathbf{\Omega},t) \approx \sum_{\nu=1}^{\mathcal{M}} M_\nu(\mathbf{\Omega})\Psi_\nu(\mathbf{r},t), \tag{6.5}$$

where $\mathbf{\Psi}(\mathbf{r},t) = (\Psi_1(\mathbf{r},t)\ \Psi_2(\mathbf{r},t)\ ...\ \Psi_{\mathcal{M}}(\mathbf{r},t))^T$ represents a vector of $\mathcal{M}$ angular moments and $M_\nu(\mathbf{\Omega})$ represents the $\mathcal{M}$ angular basis functions. The approximation, $\mathbf{S}(\mathbf{r},t)$, for the source function $s(\mathbf{r},\mathbf{\Omega},t)$, is also expanded using a discrete series of finite element basis functions:

$$s(\mathbf{r},\mathbf{\Omega},t) \approx \sum_{\nu=1}^{\mathcal{M}} M_\nu(\mathbf{\Omega})S_\nu(\mathbf{r},t), \tag{6.6}$$

where $\mathbf{S}(\mathbf{r},t) = (S_1(\mathbf{r},t)\ S_2(\mathbf{r},t)\ ...\ S_{\mathcal{M}}(\mathbf{r},t))^T$. Scattering collisions are often assumed to have rotational symmetry. The differential scattering coefficient, $\mu_s(\mathbf{r},\mathbf{\Omega}^{'}\rightarrow\mathbf{\Omega})$, therefore depends on the change in particle direction, μ_0. Consequently, $\mu_s(\mathbf{r},\mu_0)$ may be

expanded as a series of orthogonal Legendre Polynomials [112]. Following expansion, Legendre addition theorem [112] is used to expand each orthogonal Legendre polynomial as a series of directionally dependent spherical harmonics [32, 34, 59, 112, 109]. The angular discretisation of the scattering/removal operator, $\mathcal{H}$, thus yields the diagonal matrix (for spherical harmonics):

$$\begin{aligned}\mathbf{H}_{\rho\nu}(\mathbf{r}) &= \int_{\mathbf{\Omega}} (\mu_a(\mathbf{r}) + \mu_s(\mathbf{r})) M_\rho(\mathbf{\Omega}) M_\nu(\mathbf{\Omega}) d\mathbf{\Omega} \\ &- \Big[\sum_{l=0}^{L} \mu_{sl}(\mathbf{r}) \alpha^c_{\rho,l,0} \alpha^c_{n,l,0} + 2 \sum_{l=1}^{L} \mu_{sl}(\mathbf{r}) \sum_{m=1}^{l} \alpha^c_{\rho,l,m} \alpha^c_{\nu,l,m} + \alpha^s_{\rho,l,m} \alpha^s_{\nu,l,m}\Big],\end{aligned} \tag{6.7}$$

The scattering moments, $\mu_{sl}(\mathbf{r})$ have been truncated to order L. The coefficients $\alpha^c_{\rho,l,m}$ and $\alpha^s_{\rho,l,m}$ are each defined:

$$\alpha^c_{\rho,l,m} = \int_{\Omega} M_\rho(\mathbf{\Omega}) Y^c_{l,m}(\mathbf{\Omega}) d\mathbf{\Omega}, \qquad \alpha^s_{\rho,l,m} = \int_{\Omega} M_\rho(\mathbf{\Omega}) Y^s_{l,m}(\mathbf{\Omega}) d\mathbf{\Omega}. \tag{6.8}$$

Furthermore, $Y^c_{l,m}(\mathbf{\Omega})$, and $Y^s_{l,m}(\mathbf{\Omega})$ represent the real and complex parts of the spherical harmonic function, $Y_{l,m}(\mathbf{\Omega})$. If $P_{l,m}(\mu)$ denotes associated Legendre polynomials:

$$Y_{l,m}(\mathbf{\Omega}) = \begin{cases} Y^c_{l,m}(\mathbf{\Omega}) = \sqrt{\dfrac{(2l+1)(l-m)!}{(l+m)!}}\, P_{l,m}(\mu)\, cos(m\omega), \;\; 0 \leq m \leq l, \\ Y^s_{l,m}(\mathbf{\Omega}) = \sqrt{\dfrac{(2l+1)(l-m)!}{(l+m)!}}\, P_{l,m}(\mu)\, sin(m\omega), \;\; 0 \leq m \leq l. \end{cases}$$

The final form of the angularly discretised Boltzmann transport equation is a multi-dimensional differential difference equation representing a symmetric, hyperbolic coupled system:

$$\frac{1}{v}\mathbf{A_t}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial t} + \mathbf{A_x}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial x} + \mathbf{A_y}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial y} + \mathbf{A_z}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial z} + \mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r},t) - \mathbf{S}(\mathbf{r},t) = 0. \tag{6.9}$$

The three $\mathcal{M} \times \mathcal{M}$ angular Jacobian matrices, $\mathbf{A_x}, \mathbf{A_y}, \mathbf{A_z}$ and the $\mathcal{M} \times \mathcal{M}$ angular mass matrix, $\mathbf{A_t}$ are each defined:

$$\mathbf{A}_{t_{\rho\nu}} = \int_{\Omega} M_\rho(\mathbf{\Omega}) M_\nu(\mathbf{\Omega}) d\mathbf{\Omega} = \mathbf{I}, \quad \mathbf{A}_{k_{\rho\nu}} = \int_{\Omega} M_\rho(\mathbf{\Omega}) \Omega_k M_\nu(\mathbf{\Omega}) d\mathbf{\Omega}, \tag{6.10}$$

where $\mathbf{A}_k$ is the vector of angular Jacobian matrices $\mathbf{A}_k = (\mathbf{A}_x\ \mathbf{A}_y\ \mathbf{A}_z)^T$ and $k = x, y, z$. The incremental solid angle $d\mathbf{\Omega} = d\omega d\mu$. The co-ordinate, μ is the co-latitude from the z-axis and ω represents the azimuthal angle. If x, y and z are orthogonal spatial co-ordinates, the directional cosines of $\mathbf{\Omega}$ with respect to x, y and z are defined in terms of Ω_x, Ω_y, and Ωz, in which $\Omega_x = cos(\theta) = \mu$; $\Omega_y = (1-\mu^2)^{\frac{1}{2}}\cos(\omega)$; and $\Omega_z = (1-\mu^2)^{\frac{1}{2}}\sin(\omega)$.

A Streamline Upwind Petrov-Galerkin (SUPG) formulation [58, 59] is used to spatially and temporally discretise the time-dependent Boltzmann transport equation. Prior to discretisation, the angularly discretised Boltzmann transport equation is pre-multiplied by a space-time SUPG term to yield:

$$(\mathbf{I} - \mathbf{A}_{xt} \cdot \nabla_{xt}\mathbf{P}) \left(\frac{1}{v}\frac{\partial \mathbf{\Psi}(\mathbf{r},t)}{\partial t} + \mathbf{A} \cdot \nabla \mathbf{\Psi}(\mathbf{r},t) + \mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r},t) - \mathbf{S}(\mathbf{r},t) \right) = 0, \tag{6.11}$$

where $\mathbf{P}$ represents a stabilisation matrix, $\mathbf{A} = (\mathbf{A}_x\ \mathbf{A}_y\ \mathbf{A}_z)^T$, $\mathbf{A}_{xt} = (\frac{1}{v}\mathbf{I}\ \mathbf{A})^T$, and $\nabla_{xt} = (\mathbf{I}\frac{\partial}{\partial t}\ \nabla)^T$. $\mathbf{H}(\mathbf{r}) = \mathbf{H}$. The stabilisation matrix is a diagonal matrix of spherical harmonics that determines the amount of weighting assigned to a given finite element discretisation. A matrix operator, $g(h\mathbf{\Lambda}_{H_{\nu\nu}})$, is used to simplify the description of stabilisation matrices. A new stabilisation matrix $\mathbf{P}$ is defined that is differentiable for inverse problems. The scattering/removal operator, $\mathbf{H}$, represents a positive, semi-definite matrix

of non-negative eigenvalues and is defined in terms of a matrix of right eigenvectors, $\mathbf{R}_H$, a matrix of left eigenvectors, $\mathbf{R}_H^{-1}$ and a corresponding diagonal matrix of eigenvalues of $\mathbf{\Lambda}_H$. In other words, $\mathbf{H} = \mathbf{R}_H\mathbf{\Lambda}_H\mathbf{R}_H^{-1}$. When using spherical harmonic basis functions, $\mathbf{H}$ is a diagonal matrix and therefore $\mathbf{R}_H$ is the identity matrix. The matrix of modified eigenvalues is defined as:

$$\mathbf{G}_{\rho\nu}(\mathbf{\Lambda}_H) = \begin{cases} g(\mathbf{\Lambda}_{H_{\nu\nu}}) & if \;\; \rho = \nu; \\ 0 & otherwise, \end{cases} \tag{6.12}$$

in which, $g(\mathbf{\Lambda}_{H_{\nu\nu}}) = e^{-0.37h\mathbf{\Lambda}_{H_{\nu\nu}}}$. The stabilisation matrix is subsequently defined as $\mathbf{P} = h\mathbf{R}_H\mathbf{G}(\mathbf{\Lambda}_H)\mathbf{R}_H^{-1}$, in which h is the spatially varying characteristic element size. The stabilisation matrix, provides an appropriate amount of dissipation and therefore represents an optimal method to use in both optically thick and optically thin media. If the second set of brackets on the left hand side of equation 6.11 is defined as the residual. Equation 6.11 is recast:

$$(\mathbf{I} - \mathbf{A}_{xt} \cdot \nabla_{xt}\mathbf{P})\mathscr{R}(\mathbf{\Psi}(\mathbf{r}, t)) = 0. \tag{6.13}$$

The spatial and temporal dependence of the angularly discretised flux, source and scattering/removal term are discretised in space and time using a finite element approximation:

$$\begin{aligned} \boldsymbol{\psi}(\mathbf{r}, t) \approx \mathbf{\Psi}(\mathbf{r}, t) &= \sum_{\tau=1}^{\mathcal{T}} \sum_{j=1}^{\mathcal{N}_t} \mathbf{N}_j^\tau(\mathbf{r}, t)\mathbf{\Psi}_j^\tau = \sum_{k=1}^{\mathcal{N}} \mathbf{N}_k(\mathbf{r}, t)\mathbf{\Psi}_k, \\ \mathbf{s}(\mathbf{r}, t) \approx \mathbf{S}(\mathbf{r}, t) &= \sum_{\tau=1}^{\mathcal{T}} \sum_{j=1}^{\mathcal{N}_s} \mathbf{N}_{s_j}^\tau(\mathbf{r}, t)\mathbf{S}_j^\tau = \sum_{k=1}^{\mathcal{N}} \mathbf{N}_k(\mathbf{r}, t)\mathbf{S}_k, \\ \boldsymbol{\mathcal{H}}(\mathbf{r}) \approx \mathbf{H}(\mathbf{r}) &= \sum_{j=1}^{\mathcal{N}_H} \mathbf{N}_{H_j}(\mathbf{r})\mathbf{H}_j, \end{aligned}$$

in which $\mathbf{N}_k = \mathbf{N}_k(\mathbf{r}, t)$ is the amalgamation of all basis functions that are discontinuous in time, $\mathbf{N}_j^\tau(\mathbf{r}, t)$. τ represents the time level, and $\mathcal{T}$, represents the total number

of time levels. $\mathbf{N}_j^\tau(\mathbf{r},t)$, $\mathbf{N}_{s_j}^\tau(\mathbf{r},t)$, and $\mathbf{N}_{H_j}(\mathbf{r})$ are each $(\mathcal{M}\times\mathcal{M})$ diagonal matrices containing finite element basis functions associated with the angular flux, sources and scattering/removal terms respectively. In this work $\mathbf{N}_j^\tau(\mathbf{r},t)$ and $\mathbf{N}_{s_j}^\tau(\mathbf{r},t)$ have a bi-linear representation in two dimensional space and a discontinuous finite element variation in time. $\mathcal{N}_t$, $\mathcal{N}_s$ and $\mathcal{N}_H$ represent the number of space-time finite element basis functions associated with the angular flux, sources and materials (in our examples $\mathcal{N}_t=\mathcal{N}_s=2\mathcal{N}_H$ and $\mathcal{N}=\mathcal{T}\mathcal{N}_t$) and $\mathcal{N}$ represents the total number of basis functions over all time steps associated with the angular flux and source terms. $\mathbf{\Psi}_j^\tau$ and $\mathbf{S}_j^\tau$ are vectors, of size $\mathcal{M}$, containing the moments of the angularly discretised flux and source for time level, τ. $\mathbf{H}_j$ is a matrix containing the scattering/removal terms at node j. A Petrov-Galerkin method is used to reduce the single Boltzmann transport equation into a series of unknowns. Equation 6.11 is multiplied by a further matrix of finite element basis functions, $\mathbf{N}_i^\tau(\mathbf{r},t)$, before the expression is integrated over the space time slab, $V\times\Delta T^\tau$. Furthermore, Green's theorem [59] is used to convert the volume integrals representing the angularly discretised streaming terms into surface terms. This results in the expression for time step, τ:

$$
\begin{aligned}
&- \int_{\Delta T^\tau}\int_V \left(\mathbf{A}_{xt}\cdot\nabla_{xt}\mathbf{N}_i^\tau(\mathbf{r},t)\right)\mathbf{\Psi}(\mathbf{r},t)dVdt \\
&+ \int_{\Delta T^\tau}\int_V \mathbf{N}_i^\tau(\mathbf{r},t)\left(\mathbf{H}(\mathbf{r})\mathbf{\Psi}(\mathbf{r},t)-\mathbf{S}(\mathbf{r},t)\right)dVdt \\
&+ \int_{\Delta T^\tau}\int_V \left(\mathbf{A}_{xt}\cdot\nabla_{xt}\mathbf{N}_i^\tau(\mathbf{r},t)\right)\mathbf{P}\mathscr{R}dVdt + \int_{\Gamma_t^\tau}\int_V \mathbf{N}_i^\tau(\mathbf{r},t)\left(\frac{1}{v}\mathbf{I}\,n_t\right)\mathbf{\Psi}(\mathbf{r},t)dVd\Gamma_t^\tau \\
&- \int_{\Gamma_t^\tau}\int_V \mathbf{N}_i^\tau(\mathbf{r},t)\left(\frac{1}{v}\mathbf{I}\,n_t\right)\mathbf{P}\mathscr{R}dVd\Gamma_t^\tau + \int_{\Delta T^\tau}\int_{\Gamma_x} \mathbf{N}_i^\tau(\mathbf{r},t)\left(\mathbf{A}\cdot\mathbf{n}\right)\mathbf{\Psi}(\mathbf{r},t)d\Gamma_x dt \\
&- \int_{\Delta T^\tau}\int_{\Gamma_x} \mathbf{N}_i^\tau(\mathbf{r},t)\left(\mathbf{A}\cdot\mathbf{n}\right)\mathbf{P}\,\mathscr{R}\,d\Gamma_x dt = 0,
\end{aligned}
\tag{6.14}
$$

in which ΔT^τ represents the time domain of the time level τ and, Γ_t^τ, its boundary. $\mathbf{n}=(n_x\mathbf{I}\;\; n_y\mathbf{I}\;\; n_z\mathbf{I}\;\; n_t\mathbf{I})^T$ is the normal to the boundary, $(\Gamma_x\cup\Gamma_t^\tau)$, of the solution domain, $V\times\Delta T^\tau$. All terms using $\mathscr{R}$ and integrated on the boundary of the space-time domain are

assumed zero and contribute to the boundary conditions of the method. n_t is the normal component to the time boundary, Γ_t^τ, of the time domain ΔT^τ and $\mathbf{I}$ is an $(M \times M)$ identity matrix. The finite element basis function $\mathbf{N}_i^\tau(\mathbf{r}, t)$ represents a diagonal matrix:

$$\mathbf{N_i}^\tau(\mathbf{r}, t) = \begin{bmatrix} N_i^\tau(\mathbf{r}, t) & 0 & . & . & 0 \\ 0 & N_i^\tau(\mathbf{r}, t) & 0 & . & . \\ . & 0 & N_i^\tau(\mathbf{r}, t) & 0 & . \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & N_i^\tau(\mathbf{r}, t) \end{bmatrix}, \tag{6.15}$$

where $i = (1, 2, ..., \mathcal{N}_t)$, and $\mathbf{N}_{si}^\tau(\mathbf{r}, t) = \mathbf{N}_i^\tau(\mathbf{r}, t)$. The matrix $\mathbf{N}_{H_j} = \mathbf{N}_{H_j}(\mathbf{r})$ is similarly defined but has no time dependence and therefore just contains spatial finite element basis functions (which are two dimensional in the examples). A full description of the spatial and temporal discretisation of the Boltzmann transport equation using Streamline Upwinded Petrov-Galerkin methods is described in [59, 61, 129]. [59, 61, 129] also describes how estimations of the angular flux at cell faces and domain boundaries are determined through a method of mapping the angular moments to Riemann variables [58, 61, 62, 129]. Following discretisation, the direct problem is defined according to the matrix equation:

$$\mathbf{E}\boldsymbol{\Psi}_s = \mathbf{b}_s. \tag{6.16}$$

A total of $\mathcal{S}$ source problems must be solved. Each source problem results in a solution vector for the angular flux $\boldsymbol{\Psi}_s = (\boldsymbol{\Psi}_{\mathbf{s}\,1\,1}\ \boldsymbol{\Psi}_{\mathbf{s}\,1\,2}\ ...\ \boldsymbol{\Psi}_{\mathbf{s}\,\mathcal{N}\mathcal{M}})^T$, in which $s = (1, 2, ..., \mathcal{S})$. By solving the discrete linear system described in equation 6.16, a solution for the angular flux at all points along the finite element mesh is found. The GMRES iterative colver is used to solve the discrete linear system [164]. The GMRES solver incorporates block forward and backward Gauss Siedel preconditioning and, for a large number of angular moments, an angular multigrid preconditioner. Each block is an $\mathcal{M} \times \mathcal{M}$ submatrix associated with each node of the finite element mesh. Matrix $\mathbf{E}$ depends

on the discretisation method that is used and has a block sparse structure and vector, $\mathbf{b}_s = (\mathbf{b_s}_{1\,1}\ \mathbf{b_s}_{1\,2}\ ...\ \mathbf{b_s}_{\mathcal{N}\mathcal{M}})^T$, represents discretised source contributions.

6.2.2 The Finite Element Time Discretisation

The advection term in equation 6.14 was integrated using Green's theorem. This ensures that the method is conservative even when reduced or inexact quadrature is used. Incoming boundary conditions, $\mathbf{\Psi}_{in}(\mathbf{r},t)$, on $\mathbf{\Psi}(\mathbf{r},t)$ are thus applied on the surface integral:

$$\begin{aligned}\int_{\Delta T^\tau}\int_{\Gamma_x}\mathbf{N}_i^\tau(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}(\mathbf{r},t)d\Gamma_x dt &= \int_{\Delta T^\tau}\int_{\Gamma_{x_{in}}}\mathbf{N}_i^\tau(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}_{in}(\mathbf{r},t)d\Gamma_x \\ &+ \int_{\Delta T^\tau}\int_{\Gamma_{x_{out}}}\mathbf{N}_i^\tau(\mathbf{r},t)\,(\mathbf{A}\cdot\mathbf{n})\,\mathbf{\Psi}(\mathbf{r},t)d\Gamma_x,\end{aligned} \tag{6.17}$$

in which $\Gamma_{x_{in}}$ represents the incoming flux boundary and $\Gamma_{x_{out}}$ represents the outgoing flux boundary. The boundary conditions associated with the incoming and outgoing flux are approximated using a Riemann approach [58, 61, 62, 129]. The incoming flux is a specified incoming boundary condition and the outgoing flux is taken from the solution internal to the domain, but on the domain boundary, Γ_x. The global matrix, $\mathbf{E}$, described in equation 6.16, has the structure:

$$\mathbf{E} = \begin{bmatrix} \mathbf{Q}_{11} & 0 & . & . & 0 \\ \mathbf{U}_{21} & \mathbf{Q}_{22} & 0 & . & . \\ 0 & \mathbf{U}_{32} & \mathbf{Q}_{33} & 0 & . \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & 0 & \mathbf{U}_{\tau\,\tau-1} & \mathbf{Q}_{\tau\tau} \end{bmatrix}, \tag{6.18}$$

in which, the term $\mathbf{Q}_{\tau\tau}$ is defined:

$$
\begin{aligned}
\mathbf{Q}_{\tau\ \tau_{ij}} = \; & - \int_{\Delta T^\tau} \int_V \left(\mathbf{A}_{xt} \cdot \nabla_{xt} \mathbf{N}_i^\tau(\mathbf{r},t)\right) \mathbf{N}_j^\tau(\mathbf{r},t) dV dt \\
& + \int_{\Delta T^\tau} \int_V \mathbf{N}_i^\tau(\mathbf{r},t) \mathbf{H}(\mathbf{r}) \mathbf{N}_j^\tau(\mathbf{r},t) dV dt \\
& + \int_{\Delta T^\tau} \int_V \left(\mathbf{A}_{xt} \cdot \nabla_{xt} \mathbf{N}_i^\tau(\mathbf{r},t)\right) \mathbf{P} \left(\mathbf{A}_{xt} \cdot \nabla_{xt} \mathbf{N}_j^\tau(\mathbf{r},t) + \mathbf{H}(\mathbf{r}) \mathbf{N}_j^\tau(\mathbf{r},t)\right) dV dt \\
& + \int_{\Gamma^\tau_{t_{out}}} \int_V \mathbf{N}_i^\tau(\mathbf{r},t) \left(\frac{1}{v}\mathbf{I}\, n_t\right) \mathbf{N}_j^\tau(\mathbf{r},t) dV d\Gamma_t^\tau \\
& + \int_{\Delta T^\tau} \int_{\Gamma_x} \mathbf{N}_i^\tau(\mathbf{r},t) \left(\mathbf{A} \cdot \mathbf{n}\right) \mathbf{N}_j^\tau(\mathbf{r},t) d\Gamma_x dt. \qquad (6.19)
\end{aligned}
$$

$\mathbf{U}_{\tau\ \tau-1}$ is independent of the material properties and therefore does not need to be differentiated. $\mathbf{U}_{\tau\ \tau-1}$ is defined:

$$
\mathbf{U}_{\tau\ \tau-1} = \int_{\Gamma^\tau_{t_{in}}} \int_V \mathbf{N}_i^\tau(\mathbf{r},t) \left(\frac{1}{v}\mathbf{I}\ \ n_t\right) \mathbf{N}_j^\tau(\mathbf{r},t)\ dV\ d\Gamma_t^\tau \quad for \quad n_t < 0, \qquad (6.20)
$$

The last surface term involving $\mathscr{R}$ in equation 6.14 is assumed to be equal to zero. That is, the residual for both the outgoing and incoming flux is assumed to be zero as a natural boundary condition:

$$
\int_{\Delta T^\tau} \int_{\Gamma_x} \mathbf{N}_i^\tau(\mathbf{r},t) \left(\mathbf{A} \cdot \mathbf{n}\right) \mathbf{P}\, \mathscr{R}\ d\Gamma_x dt = 0. \qquad (6.21)
$$

and similarly on the time boundary:

$$
\int_{\Gamma_t^\tau} \int_V \mathbf{N}_i^\tau(\mathbf{r},t) \left(\mathbf{A} \cdot \mathbf{n}\right) \mathbf{P}\, \mathscr{R}\ dV d\Gamma_t^\tau = 0. \qquad (6.22)
$$

This enhances the stabilisation of the time discretisation. For each of the model geometries considered in this chapter, velocity is set arbitrarily to 1 and a Discontinuous Galerkin

time step of $\Delta t = 5$ is used. The Discontinuous Galerkin time stepping method has the ability to use large time steps while still capturing the essential features of the time dependent signals with a high order of accuracy. The accuracy in time is second order, at the end of a time level, and third order at the beginning of the time level. Another desirable property of the chosen time discretisation is its ability to accurately resolve low frequencies while filtering out unresolved high frequencies.

6.2.3 The Treatment of Sources and Detectors

In operator form, the forward model describes the angular flux, ψ_s, that results from a source placed in region, $\mathbf{r}$ at a boundary location, $\mathbf{r}_s$:

$$\left[\frac{1}{v}\frac{\partial}{\partial t} + \mathbf{\Omega}\cdot\boldsymbol{\nabla} + \mathcal{H}\right]\psi_s = C_s\delta(\mathbf{r}-\mathbf{r}_s)\delta(t). \tag{6.23}$$

Equation 6.23 describes a discretised series of $\mathcal{S}$ source problems, where $s = (1, 2, ..., \mathcal{S})$. The source is a Dirac-delta function in time, $\delta(t)$, and space, $\delta(\mathbf{r}-\mathbf{r}_s)$ and is centred at position $\mathbf{r}_s$ of the domain $\mathbf{r}$ at time t. The magnitude of the source strength is given by C_s. ψ_s represents the angular flux owing to the source number, s, while the scattering/removal operator, $\mathcal{H}$, contains the spatially varying material properties. Each source problem yields an angular flux solution at all points in the solution domain. If the source position is located at a node then the source strength and angular flux are equal to the finite element basis function centred on that node multiplied by a scalar (to scale the strength):

$$\mathbf{b}_{s_i} = \int_T\int_V \mathbf{N}_i(\mathbf{r},t)\mathbf{S}_s dV dt, \tag{6.24}$$

in which, $T = (V\times\Delta T^1\ V\times\Delta T^2\ ...V\times\Delta T^\tau)^T$ represents the time domain of interest and $\mathbf{N}_i = \mathbf{N}_i(\mathbf{r},t)$ is the amalgamation of all Discontinuous Galerkin basis functions $\mathbf{N}_i^\tau(\mathbf{r},t)$. $\mathbf{S}_s$ is a finite element approximation of the source. If the source is located inside an element it must be distributed to the nearest node associated with that element and possibly nodes associated with neighbouring elements. A Gaussian function may be

used to interpolate photons from a source on to adjacent nodes. Thus, for the i^{th} nodal position, $\mathbf{x}_i$, and the s^{th} source at position $\mathbf{r}_s$, $\mathbf{S}_{si}$ is defined at each node according to:

$$\mathbf{S}_{si} = \frac{1}{n_s} exp\left[-\frac{(\mathbf{x}_i - \mathbf{r}_s)^2}{l_s^2}\right] \mathcal{U}(t), \tag{6.25}$$

where $\mathcal{U}(t) = 1$ if $t \leq \Delta t$ and is 0 otherwise. Δt represents the width of the time element and $t = 0$ at the start of the time domain. The source has been chosen to approximate a delta function in time. In other words the source is switched on only during the first Discontinuous Galerkin time step. Only the first moment source vectors are set to a non-zero value as this corresponds to an isotropic source when spherical harmonic basis functions are used. The angular flux received at a detector location is also calculated using a Gaussian function to interpolate the flux from any adjacent nodes on the surface of the detector. Detectors are distributed to the nearest nodes in a similar way as sources are distributed (i.e. according to equations 6.24, and 6.25).

6.3 The Functional

The functional that is optimised to find a solution to the inverse problem is given by:

$$F = F_d + F_r + F_v. \tag{6.26}$$

In which, F_d represents the data misfit between observed measurements and predicted data and F_r and F_v represent regularisation terms that encourage certain types of model. The model covariance information contained in the regularisation term applies smoothness and anisotropy constraints affecting a homogeneous isotropic model.

6.3.1 The Data Misfit

The error functional is a quantification of the data misfit between model predictions and measured data:

$$F_d = \frac{1}{2}\sum_{s=1}^{\mathcal{S}}\sum_{r=1}^{\mathcal{R}}\sum_{\tau=1}^{\mathcal{T}}\sum_{\mu=1}^{\mathcal{M}}\tilde{w}_s^r\left(\hat{\Psi}_{s\,\tau\,\mu} - d_{s\,\tau\,\mu}^r\right)^2, \tag{6.27}$$

in which $\hat{\Psi}_{s\,\tau\,\mu}^r = \sum_{i=1}^{\mathcal{N}}\tilde{W}_{s\,i}^r\Psi_{s\,\tau\,\mu\,i}$. $\tilde{W}_{s\,i}^r$ is chosen such that $\sum_{i=1}^{\mathcal{N}}\tilde{W}_{s\,i}^r = 1$ and may be used to estimate the angular flux at a detector by interpolating the flux from nearby nodes. The term contained in brackets on the left hand side of equation 6.27 is redefined:

$$\hat{\Psi}_{s\,\tau\,\mu}^r - d_{s\,\tau\,\mu}^r = \sum_{i=1}^{\mathcal{N}}\tilde{W}_{s\,i}^r(\Psi_{s\,\tau\,\mu\,i}^r - d_{s\,\tau\,\mu}^r). \tag{6.28}$$

Thus we can write the data mismatch functional as:

$$F_d = \frac{1}{2}\sum_{s=1}^{\mathcal{S}}\sum_{r=1}^{\mathcal{R}}\sum_{\tau=1}^{\mathcal{T}}\sum_{\mu=1}^{\mathcal{M}}\sum_{i=1}^{\mathcal{N}}W_{s\,i}^r\left(\Psi_{s\,\tau\,\mu\,i} - d_{s\,\tau\,\mu}^r\right)^2. \tag{6.29}$$

Equation 6.29 represents the weighted sum of the squared differences between observed data, $d_{s\,\tau\,\mu}^r$, and data predicted from a given material property distribution, $\Psi_{s\,\tau\,\mu\,i}$, for all $\mathcal{S}$ sources and $\mathcal{R}$ detectors. The summation is also performed over each node, $i = (1, 2, ..., \mathcal{N})$, for each time step, $\tau = (1, 2, ..., \mathcal{T})$ and for each degree of freedom, $\mu = (1, 2, ..., \mathcal{M})$. The weighting factor, $W_{s\,i}^r$, distributes both the sources and detectors to the nearest nodes using a Gaussian approximation. For the i^{th} node at position $\mathbf{x}_i$, the r^{th} detector at position, $\mathbf{r}_s^r$, and the s^{th} source experiment, the weighting contribution is defined:

$$W_{s\,i}^r = \frac{w_s^r}{n_s^r}exp\left[-\frac{\left(\mathbf{x}_i - \mathbf{r}_s^r\right)^2}{l_r^2}\right]. \tag{6.30}$$

Where l_r is the length over which the interpolation is performed and n_s^r is a normalisation

coefficient. Equation 6.30 provides a mechanism for gathering the flux from local nodes when the detector is placed between nodes. If the size of the Gaussian width, l_r, is small compared to the element size then equation 6.30 assumes that the detector is re-positioned at the nearest node. In all examples in this thesis the detector positions are centred on nodes and l_r is small relative to the neighbouring element sizes. Thus the use of equation 6.30 takes the flux at the node in question and compares it with the idea detector response in equation 6.29. Therefore, $W_s^{r_k} = w_s^r/n_s^r$ for a single node $k = j$ (j is the node at which the detector is placed) otherwise, $W_s^{r_k} = 0$. The weighting function typically has three contributions. In the first two contributions w_s^r is used to weight the datum associated with each source-detector pair according to 1) the confidence in the datum and 2) for preferential weighting of datum. When using w_s^r for preferential weighting, large weights are typically assigned to small data-values. The final contribution in the weight, $W_{s\,i}^r$, arises from the way in which the detectors are treated. The detectors are distributed to the nearest node using a Gaussian distribution in the same way as sources are distributed. The weight function, w_s^r, is defined:

$$w_s^r = \begin{cases} 0 & if \quad ((x_r - x_s)^2 + (y_r - y_s)^2 + (z_r - z_s)^2) < \kappa^2; \\ \left(\dfrac{d_{max_s}}{\phi_1 d_{max_s} + \phi_2 d^r_{max_s}} \right)^2 & otherwise, \end{cases}$$

where $\kappa^2 = (l_s + l_r)^2$ and l_s and l_r represent the length over which the interpolation is performed. x_s, x_r, y_s, y_r, z_s and z_r represent the x, y and z position coordinates for the sources and detectors respectively. ϕ_1 and ϕ_2 are carefully chosen scalar constants that are used to adjust weights so that data at detectors with relatively small scalar fluxes can be given higher priority. κ is chosen to avoid putting excessive emphasis on short range inversion information obtained by placing sources and detectors close to one another. d_{max_s} is associated with the maximum value of the observed scalar flux for all detectors and for all time:

$$d_{max_s} = \max_r \{d^r_{max_s}\}, \tag{6.31}$$

and $d^r_{max_s}$ is the maximum value of the observed scalar flux at detector, r, for source

experiment, s, over all time. One of the strengths of the algorithm presented is the independence of the source and detector locations and the finite element mesh employed. In matrix form, the functional in equation 6.29 is defined:

$$F_d = \frac{1}{2}(\hat{\mathbf{\Psi}} - \mathbf{d})^{\mathbf{T}}\mathbf{W}(\hat{\mathbf{\Psi}} - \mathbf{d}), \tag{6.32}$$

in which, $\hat{\mathbf{\Psi}}^T = (\mathbf{\Psi}_1^T \quad \mathbf{\Psi}_2^T \quad \cdots \quad \mathbf{\Psi}_{\mathcal{S}}^T, \quad \cdots \quad , \mathbf{\Psi}_1^T \quad \mathbf{\Psi}_2^T \quad \cdots \quad \mathbf{\Psi}_{\mathcal{S}}^T)$, and $\mathbf{d}^T = (\mathbf{d}_1^{1T} \quad \mathbf{d}_2^{1T} \quad \cdots \quad \mathbf{d}_{\mathcal{S}}^{1T}, \quad \cdots \quad , \mathbf{d}_1^{\mathcal{R}T} \quad \mathbf{d}_2^{\mathcal{R}T} \quad \cdots \quad \mathbf{d}_{\mathcal{S}}^{\mathcal{R}T})$. Both the predicted data, $\mathbf{\Psi_s}$, and the observed data, $\mathbf{d}_s^r$, vectors are the same length as the number of nodes. The data covariance matrix $\mathbf{W}^{-1}$ is diagonal and contains the preferential weighting, w_s^r.

6.3.2 Model Covariance - Spatial Regularisation

The regularisation term defined in equation 6.26 can be used to constrain the inversion model and guide the search process in a specific direction. F_r is re-defined:

$$F_r = \frac{1}{2}\gamma_a \int_V \nabla^T m_a k \, \nabla m_a \, dV + \frac{1}{2}\gamma_{s0} \int_V \nabla^T m_{s0} k \, \nabla m_{s0} \, dV = \frac{1}{2}\mathbf{m}^T \mathbf{K} \mathbf{m}, \tag{6.33}$$

in which, $\mathbf{K}$ is represents a diffusion tensor and, in this work, is assumed to be isotropic and equal to one. γ_a and γ_{s0} control the level of model covariance imposed for absorption and scattering contributions respectively. $\mathbf{m}_a = (m_{a1} \; m_{a2} \; ... \; m_{a\mathcal{N}_H})^T$ represents the discretised absorption material properties where -assuming m_{a_i} is the nodal value of a- $m_{a_i} = ln(\mu_{a_i})$. $\mathbf{m}_{s0} = (m_{s01} \; m_{s02} \; ... \; m_{s0\mathcal{N}_H})^T$ represents the discretised scattering material properties in which $m_{s0_i} = ln(\mu_{s0_i})$. $\mathbf{m}_a$ and $\mathbf{m}_{s0}$ are defined collectively in terms of $\mathbf{m}^T = (\mathbf{m}_a^T \; \mathbf{m}_{s0}^T)$, and are approximated using a set of finite element basis functions, N_{H_j}, such that:

$$m_a = \sum_{j=1}^{\mathcal{N}_H} N_{H_j} m_{a_j}, \qquad m_{s0} = \sum_{j=1}^{\mathcal{N}_H} N_{H_j} m_{s0_j}. \tag{6.34}$$

Material properties are considered as opposed to material cross-sections because the resulting contributions to the error functional are quadratic. The error functional is therefore easily optimised using standard least squares techniques and results in uniform material properties. In addition, material cross-sections, sometimes vary by several orders of magnitude, which is undesirable when considering model regularisation. Structure is penalised using a symmetric, positive-definite matrix, $\mathbf{K}$:

$$\mathbf{K}_{ij} = \begin{bmatrix} \gamma_a \int_V \nabla N_{H_i}\, k\, \nabla N_{H_j}\, dV & 0 \\ 0 & \gamma_{s0} \int_V \nabla N_{H_i}\, k\, \nabla N_{H_j}\, dV \end{bmatrix}. \tag{6.35}$$

In which $\mathbf{K}$ represents a diffusion matrix and is defficient in rank by 1. Therefore, $\mathbf{K}$ is not invertible and a singular value would have to be set in order to make to make $\mathbf{K}$ non-singular. k is a positive semi-definite 3×3 matrix. The applied regularisation can vary from node to node by making k a function of space. k defines a scalar product, making the integrand positive throughout the domain and therefore F_r positive. For medical imaging applications inhomogeneous smoothing is required, therefore each of the elements contained inside matrix $\mathbf{K}$ should be defined as a function of position. Large initial values of γ_a or γ_{s0} affect a smooth, homogeneous model allowing large scale structures to emerge during early iterations, while during later iterations, structurally complex models emerge as, γ_a and γ_{s0}, are relaxed.

6.3.3 Model Covariance - Step Length Damping

The penalty deviation associated with an updated model relative to the starting model is achieved by replacing the absorption material property, $\mathbf{m}_a$, with $\Delta\mathbf{m}_a = \mathbf{m}_a - \mathbf{m}_{a_{obs}}$ and the scattering material property, $\mathbf{m}_s$, with $\Delta\mathbf{m}_{s0} = \mathbf{m}_{s0} - \mathbf{m}_{s0_{obs}}$:

$$\begin{aligned} F_v &= \frac{1}{2}\lambda \int_V (m_a - m_{a_{obs}})^2 dV + \frac{1}{2}\lambda \int_V (m_{s0} - m_{s0_{obs}})^2 dV \\ &\approx \frac{1}{2}\Delta\mathbf{m}_a^T \mathbf{V}\mathbf{M}_{H_L} \Delta\mathbf{m}_a + \frac{1}{2}\Delta\mathbf{m}_{s0}^T \mathbf{V}\mathbf{M}_{H_L} \Delta\mathbf{m}_{s0}, \end{aligned} \tag{6.36}$$

in which $\Delta\mathbf{m}_a = (\Delta m_{a1}\ \Delta m_{a2}\ ...\ \Delta m_{a\mathcal{N}_H})^T$ and $\Delta\mathbf{m}_{s0} = (\Delta m_{s01}\ \Delta m_{s02}\ ...\ \Delta m_{s0\mathcal{N}_H})^T$ represent the change in the absorption and scattering material properties. The absorption and scattering material properties associated with the previous least squares iteration are given by $\mathbf{m}_{a_{obs}} = (m_{a_{obs}1}\ m_{a_{obs}2}\ ...\ m_{a_{obs}\mathcal{N}_H})^T$ and $\mathbf{m}_{s0_{obs}} = (m_{s0_{obs}1}\ m_{s0_{obs}2}\ ...\ m_{s0_{obs}\mathcal{N}_H})^T$ respectively. The diagonal lumped mass matrix, $\mathbf{M}_{H_L}$ is $\mathbf{M}_{H_{L_{ii}}} = \int_V N_{H_i} dV$, where the scalar λ controls the magnitude of the step length damping. Matrix, $\mathbf{V}_{ij} = \lambda$ if $i = j$ and is 0 otherwise.

6.4 Least Squares Inversion

The adjoint form of equation 6.16 is defined:

$$\mathbf{E}^T\mathbf{\Psi}_s^* = \frac{\partial F}{\partial \mathbf{\Psi}_s}, \tag{6.37}$$

where, $\mathbf{\Psi}_\mathbf{s}^* = (\mathbf{\Psi}_{\mathbf{s}\,1\,1}^*\ \mathbf{\Psi}_{\mathbf{s}\,1\,2}^*...\mathbf{\Psi}_{\mathbf{s}\,\mathcal{N}\mathcal{M}}^*)^T$ represents the vector of the (adjoint) control variables and F represents the error functional defined in equation 6.26. Equation 6.37 describes a series of $\mathcal{S}$ source problems, where $s = (1, 2, ..., \mathcal{S})$. The differential term on the right hand side of equation 6.37 represents a time-varying data residual, which is used as a source term for subsequent adjoint calculations. Implementing $\mathbf{E}^T$ results in a backward stepping, inverse solution in both spatial and time coordinates. Using a least squares, Levenberg-Marquardt method with step length damping the following equation is solved for the solution minimum:

$$(\mathbf{J}^T\mathbf{W}\mathbf{J} + \mathbf{K} + \lambda\mathbf{M}_{H_L})\Delta\mathbf{m} = -\mathbf{J}^T\mathbf{W}^T(\mathbf{\Psi}(m_{old}) - \mathbf{d}) - \mathbf{K}\mathbf{m}_{old}. \tag{6.38}$$

In which, the model update $\Delta\mathbf{m} = \mathbf{m}_{new} - \mathbf{m}_{old}$. The terms on the right hand side of equation 6.38 represent the gradient vector ($\mathbf{J}^T\mathbf{W}^T(\mathbf{\Psi}(m_{old}) - \mathbf{d}) - \mathbf{K}\mathbf{m}_{old}$), while the terms on the left hand side represent a modified, approximate Hessian matrix ($\mathbf{J}^T\mathbf{W}\mathbf{J} + \mathbf{K} + \lambda\mathbf{M}_{H_L}$) The matrix equation is solved using a Preconditioned Conjugate Gradient (PCG) technique, in which the matrix, $\mathbf{J}^T\mathbf{W}\mathbf{J}+\mathbf{K}+\lambda\mathbf{M}_{H_L}$, is not explicitly formed. The

PCG solver requires the formation of a preconditioning matrix, $\hat{\mathbf{B}}$. The preconditioner is usually defined in such a way that the system, $\hat{\mathbf{B}}\Delta\mathbf{m} = \mathbf{a}$, amounts to a simplified version of the original system, $\mathbf{B}\Delta\mathbf{m} = \mathbf{a}$ (defined in equation 5.43). The preconditioner, $\hat{\mathbf{B}}$, therefore represents an approximation to the approximate Hessian matrix, $\mathbf{B}$ and is defined:

$$\hat{\mathbf{B}} = \mathbf{K} + \lambda \mathbf{M}_{H_L}. \tag{6.39}$$

The preconditioner, $\hat{\mathbf{B}}$, is a non-singular, positive definite matrix containing structural constraints and step length damping terms, thus enabling model covariance to be rigorously enforced during each PCG iteration. The step length damping coefficient, λ, controls the magnitude of the model update, $\Delta\mathbf{m}$, and is adjusted according to the following heuristics: 1) λ is increased by a factor of 10 if either the PCG method is not converging well or the updated model is worse than the previous model 2) If using a larger step length damping coefficient does not decrease the functional, F, the previous step length damping coefficient is used. 3) If PCG solution of equation 6.38 is converging well and the new functional is less than the old functional, then λ is reduced by a factor of 10. If λ is chosen to be large then $\Delta\mathbf{m}$ stays small. The magnitude of the spatial regularisation, γ, influences the relative importance of the structure penalty to the error functional. γ is reduced according to, $\gamma_l = \nu\gamma_{l-1}$, where l is the iteration level. The relaxation factor ν is typically set to 0.5. Large values of γ are initially chosen to affect a smooth model and are gradually annealed downwards so that complex structures emerge in later iterations.

6.4.1 Calculations Involving the Jacobian and Hessian Matrices

Both the Jacobian matrix, $\mathbf{J}$, and the approximate Hessian matrix, $\mathbf{J}^T\mathbf{W}\mathbf{J}$, are never explicitly computed or stored. Only the matrix-vector multiplication using these two matrices is required. The Jacobian matrix, $\mathbf{J}$, illustrated in equation 6.38 is given by:

$$\mathbf{J} = \mathcal{I}^T \begin{pmatrix} \mathcal{J}_1 \\ \mathcal{J}_2 \\ .. \\ .. \\ \mathcal{J}_{\mathcal{S}} \end{pmatrix}, \tag{6.40}$$

in which $\mathcal{S}$ represents the total number of sources and $\mathcal{I}^T = (I_1, I_2, ..., I_{\mathcal{S}})$. Each I_s for $s = (1, 2, ..., \mathcal{S})$ represents an identity matrix of order $\mathcal{N}_H \times \mathcal{N}_H$. Each $\mathcal{J}_s$ contains the change in the angular flux at a given node caused by a perturbation in the material properties at another node for every source experiment, s. The Jacobian, $\mathcal{J}_s$ of the s^{th}-source experiment is approximated according to:

$$\mathcal{J}_s = \mathbf{E}^{-1}\mathbf{L}_s, \tag{6.41}$$

in which, $\mathbf{L}_s = (\mathbf{l}_{a_{s1}}\ \mathbf{l}_{a_{s2}}\ ...\ \mathbf{l}_{a_{s\mathcal{N}_H}}\ \mathbf{l}_{s0_{s1}}\ \mathbf{l}_{s0_{s2}}\ ...\ \mathbf{l}_{s0_{s\mathcal{N}_H}})$. Matrix $\mathbf{E}$ describes the direct problem in equation 6.16. The j^{th} column vector, associated with absorption material properties, $\mathbf{l}_{a_{sj}}$ is defined as:

$$\mathbf{l}_{a_{sj}} = \frac{\partial \mu_{a_j}}{\partial m_{a_j}} \left(\frac{\partial \mathbf{E}}{\partial \mu_{a_j}} \mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial \mu_{a_j}} \right). \tag{6.42}$$

The j^{th} column vector, associated with scattering material properties, $\mathbf{l}_{s0_{sj}}$ is defined as:

$$\mathbf{l}_{s0_{sj}} = \frac{\partial \mu_{s0_j}}{\partial m_{s0_j}} \left(\frac{\partial \mathbf{E}}{\partial \mu_{s0_j}} \mathbf{\Psi}_s - \frac{\partial \mathbf{b}_s}{\partial \mu_{s0_j}} \right). \tag{6.43}$$

Both an absorption coefficient and a scattering coefficient are defined for each cell. The partial derivatives of the scattering and absorption coefficients with respect to material property control variables, m_{a_j} and m_{s0_j} have been defined $\frac{\partial \mu_{a_j}}{\partial m_{a_j}} = exp(m_{a_j})$ and $\frac{\partial \mu_{s0_j}}{\partial m_{s0_j}} = exp(m_{s0_j})$ respectively, which are exact for each node and ensures positivity.

Furthermore the scattering/removal operator is a sparse diagonal matrix. The diagonal terms represent material properties.

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \mu_a + \mu_{s0} - \mu_{s0} & 0 & . & . & 0 \\ 0 & \mu_a + \mu_{s0} - \mu_{s1} & 0 & . & . \\ . & 0 & \mu_a + \mu_{s0} - \mu_{s2} & 0 & . \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & \mu_a + \mu_{s0} - \mu_{s\mathcal{M}-1} \end{bmatrix}. \tag{6.44}$$

The term contained in brackets in equation 6.42 can be formulated in terms of the gradient of the absorption material properties:

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial \mu_{a_j}} \boldsymbol{\Psi}_s \quad - \quad \frac{\partial \mathbf{b}_s}{\partial \mu_{a_j}} = \int_T \int_V \Big\{ & (\mathbf{N}_j - (\nabla_{xt} \mathbf{N}_j \cdot \mathbf{A}_{xt}) \mathbf{P}) \frac{\partial \mathbf{H}}{\partial \mu_{a_j}} \boldsymbol{\Psi}_s \\ - \quad & (\nabla_{xt} \mathbf{N}_j \cdot \mathbf{A}_{xt}) \left(\frac{\partial \mathbf{P}}{\partial \mu_{a_j}} \right) \left(\frac{1}{v} \frac{\partial \boldsymbol{\Psi}_s}{\partial t} + \mathbf{A} \cdot \nabla \boldsymbol{\Psi}_s + \mathbf{H} \boldsymbol{\Psi}_s - \mathbf{S}_s \right) \Big\} dV dt, \end{aligned} \tag{6.45}$$

and the gradient of the scattering material properties repectively:

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial \mu_{s0_j}} \boldsymbol{\Psi}_s \quad - \quad \frac{\partial \mathbf{b}_s}{\partial \mu_{s0_j}} = \int_T \int_V \Big\{ & (\mathbf{N}_j - (\nabla_{xt} \mathbf{N}_j \cdot \mathbf{A}_{xt}) \mathbf{P}) \frac{\partial \mathbf{H}}{\partial \mu_{s0_j}} \boldsymbol{\Psi}_s \\ - \quad & (\nabla_{xt} \mathbf{N}_j \cdot \mathbf{A}_{xt}) \left(\frac{\partial \mathbf{P}}{\partial \mu_{s0_j}} \right) \left(\frac{1}{v} \frac{\partial \boldsymbol{\Psi}_s}{\partial t} + \mathbf{A} \cdot \nabla \boldsymbol{\Psi}_s + \mathbf{H} \boldsymbol{\Psi}_s - \mathbf{S}_s \right) \Big\} dV dt, \end{aligned} \tag{6.46}$$

where $\mathbf{S}_s$ represent the source for source problem s. The scattering material properties incorporate both the scattering coefficient, μ_s, and the scattering moments, μ_{sl}. Since

the nodal value for μ_{a_j} or μ_{s0_j} only manifests itself in $\mathbf{E}$ in a small region of the local vicinity of node, j, $\mathbf{l}_{a_{sj}}$ and $\mathbf{l}_{s0_{sj}}$ have very few non-zero entries. Manipulations involving $\mathbf{l}_{a_{sj}}$ or $\mathbf{l}_{s0_{sj}}$ are therefore computationally efficient. Matrix-vector multiplications involving, $\mathbf{L}_s$, or its transpose, $\mathbf{L}_s^T$, are demonstrated using the vectors, $\mathbf{q} = (q_{a_1}\ q_{a_2}\ ...\ q_{a_{\mathcal{N}_H}}\ q_{s0_1}\ q_{s0_2}\ ...\ q_{s0_{\mathcal{N}_H}})^T$ and $\mathbf{v} = (v_{a_1}\ v_{a_2}\ ...\ v_{a_{\mathcal{N}_H}}\ v_{s0_1}\ v_{s0_2}\ ...\ v_{s0_{\mathcal{N}_H}})^T$ respectively, in which:

$$\mathbf{L}_s\mathbf{q} = (q_{a_1}\mathbf{l}_{a_{s1}} + q_{a_2}\mathbf{l}_{a_{s2}} + ..., q_{a_{\mathcal{N}_H}}\mathbf{l}_{a_{s\mathcal{N}_H}} + q_{s0_1}\mathbf{l}_{s0_{s1}} + q_{s0_2}\mathbf{l}_{s0_{s2}} + ..., q_{s0_{\mathcal{N}_H}}\mathbf{l}_{s0_{s\mathcal{N}_H}}), \tag{6.47}$$

$$\mathbf{L}_s^T\mathbf{v} = (v_{a_1}\mathbf{l}_{a_{s1}} + v_{a_2}\mathbf{l}_{a_{s2}} + ..., v_{a_{\mathcal{N}_H}}\mathbf{l}_{a_{s\mathcal{N}_H}} + v_{s0_1}\mathbf{l}_{s0_{s1}} + v_{s0_2}\mathbf{l}_{s0_{s2}} + ..., v_{s0_{\mathcal{N}_H}}\mathbf{l}_{s0_{s\mathcal{N}_H}})^T. \tag{6.48}$$

Multiplications involving $\mathbf{L}_s$ and $\mathbf{L}_s^T$ are required to form the approximate Hessian matrix, $\mathbf{J}^T\mathbf{W}\mathbf{J}\mathbf{q}$. The matrix vector multiplication for an arbitrary vector, $\mathbf{q}$, of length $2\mathcal{N}_H$, becomes:

$$\mathbf{J}^T\mathbf{W}\mathbf{J}\mathbf{q} = \sum_{s=1}^{\mathcal{S}} \mathbf{L}_s^T\mathbf{E}^{-T}\mathcal{W}_s\mathbf{E}^{-1}\mathbf{L}_s\mathbf{q}, \tag{6.49}$$

where $\mathcal{W}_s = \sum_{r=1}^{\mathcal{R}} \mathbf{W}_s^r$ and the diagonal matrix $\mathbf{W}_s^r$ has diagonal elements w_s^r. To obtain vector, $\mathbf{J}^T\mathbf{W}\mathbf{J}\mathbf{q}$, the solution of two matrix equations involving matrix $\mathbf{E}$ are required for each source problem.

6.5 Numerical Examples

Presented are the simultaneous reconstructions of absorption and scattering coefficients contained inside a given isotropic medium from measurements taken at the surface of the domain. Circular and square test objects containing embedded inhomogeneities with dif-

	Experiment 1		Experiments 2 and 4		Experiment 3	
Source/Detector Number	x	y	x	y	x	y
1	-2.5	0.0	1.0	0.0	0.0	3.5
2	-1.75	1.75	2.0	0.0	2.5	2.5
3	0.0	2.5	3.0	0.0	3.5	0.0
4	1.75	1.75	4.0	0.0	2.5	-2.5
5	2.5	0.0	5.0	1.0	0.0	-3.5
6	1.75	-1.75	5.0	2.0	-2.5	-2.5
7	0.0	-2.5	5.0	3.0	-3.5	0.0
8	-1.75	-1.75	5.0	4.0	-2.5	2.5
9			4.0	5.0		
10			3.0	5.0		
11			2.0	5.0		
12			1.0	5.0		
13			0.0	4.0		
14			0.0	3.0		
15			0.0	2.0		
16			0.0	1.0		

Table 6.1: The x co-ordinate and y co-ordinate source and detector locations for each of the time dependent inversion experiments described in chapter 6.

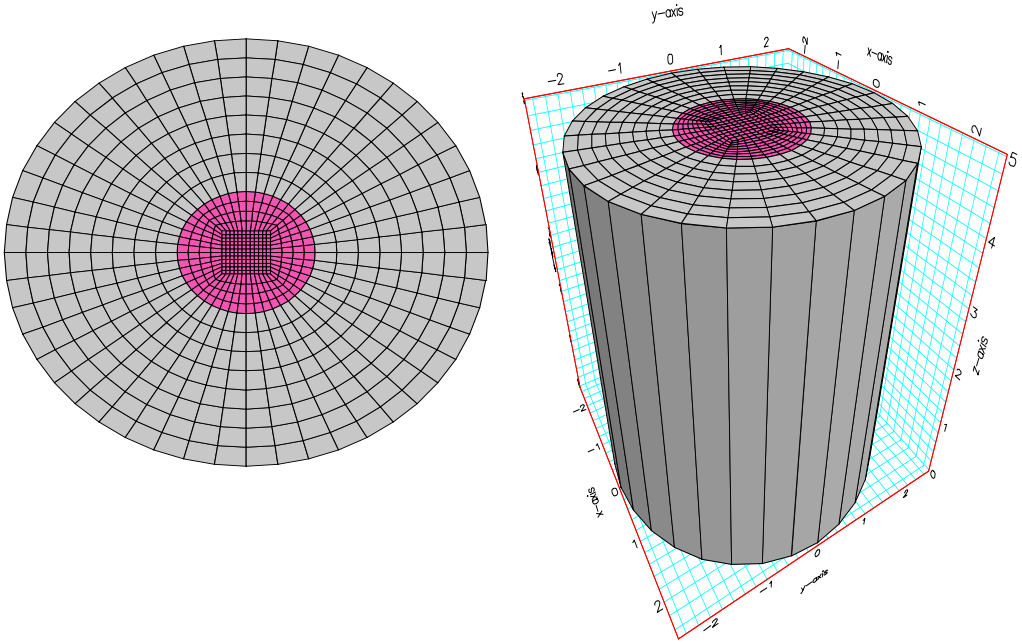


Figure 6.1: A two dimensional structured finite element circular mesh of diameter 5cm. An absorbing object of diameter 2cm is located at the centre of the domain. The time dependent problem has 1186 nodes, 1760 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps.

ferent scattering and absorption parameters are utilised as a crude representation of human tissue or a human organ. The forward (light transport) model used in the reconstruction of each of the test cases was simulated using the SUPG, finite element representation of the time dependent Boltzmann transport equation described in the theory. The inversion calculations utilise an iterative, gradient based reconstruction scheme that repeatedly calculates an update of the solution until an objective function, becomes acceptably small. The functional is minimised on a structured finite element mesh using a gradient based optimisation scheme. Highly resolved meshes are used for both forward and adjoint calculations to ensure good quality data. For each of the calculations described, a homogeneous medium with optical properties equal to those of the background medium is provided as an initial estimate. The various locations of the individual sources and detectors for each of the subsequent experiments are described in table 6.1.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.2	7.0
Inhomogenity	0.1	5.0

Table 6.2: Optical absorption and scattering coefficients for a circle 5cm in diameter and an inhomogeneity placed at the centre of that circle 2cm in diameter.

6.5.1 Experiment 1

The first simulation geometry depicts a circular domain 5cm in diameter containing an inhomogeneity situated in the centre of the circle that is 2cm in diameter (figure 6.1). The absorption and scattering coefficients of the medium and the inhomogeneity are contained in table 6.2. Eight equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward model and reconstruction model results. A P_1 angular approximation is used to obtain both the time-dependent forward and inversion model calculations along with a Levenberg-Marquardt optimisation technique. The same sized mesh is used to generate both the synthetic data and the inversion reconstruction. The mesh used for the two dimensional forward model and inverse model reconstruction had 1186 nodes, 1760 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium is chosen with optical properties that equal the background. In other words $\mu_a = 0.2$cm and $\mu_s^{'} = 7.0$cm.

Figure 6.2 illustrates scalar flux as a function of time for the first source position at the third detector location for the problem geometry described in figure 6.1. The source was placed at position co-ordinates (-2.5, 0.0) and the detector was placed at position co-ordinates (0.0, 2.5). A Discontinuous Galerkin time discretisation technique is used to discretise the time domain. Figure 6.2 depicts the exact solution detector response using 5 (solid line) and 25 (dashed line) Discontinuous Galerkin time steps and illustrates that the Discontinuous Galerkin method has the ability to use large time steps while still capturing

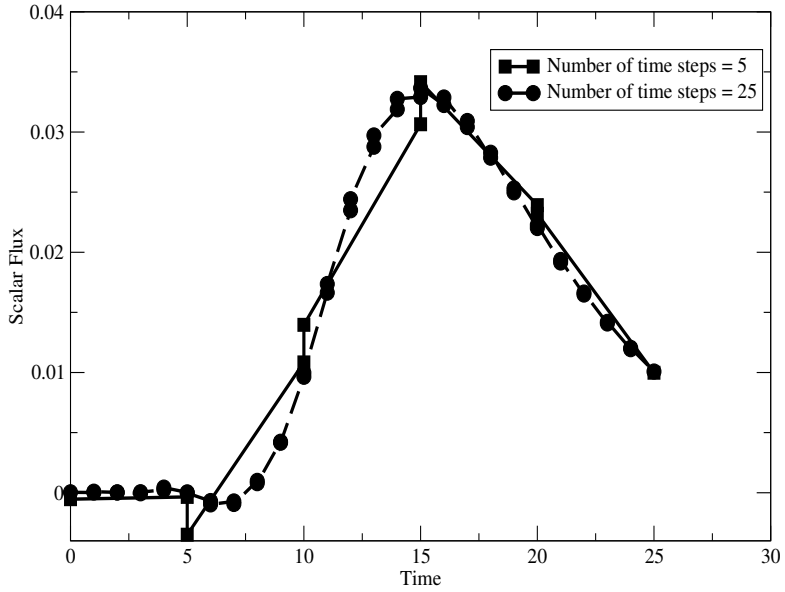
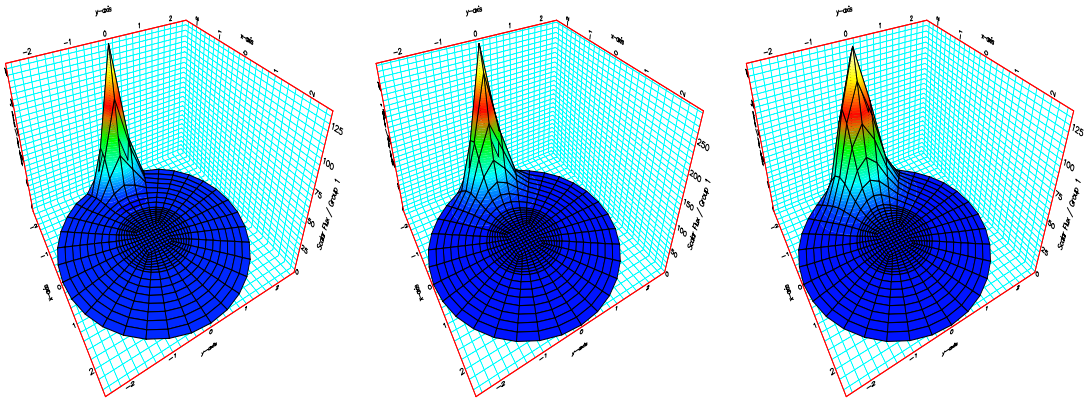


Figure 6.2: Scalar flux versus time for the first source position observed at the third detector location. The source was placed at co-ordinates (-2.5,0.0) and the detector was placed at co-ordinates (0.0 2.5). Figure 6.2 illustrates the exact solution response using 5 (solid line) and 25 (dashed line) Discontinuous Galerkin time steps respectively.

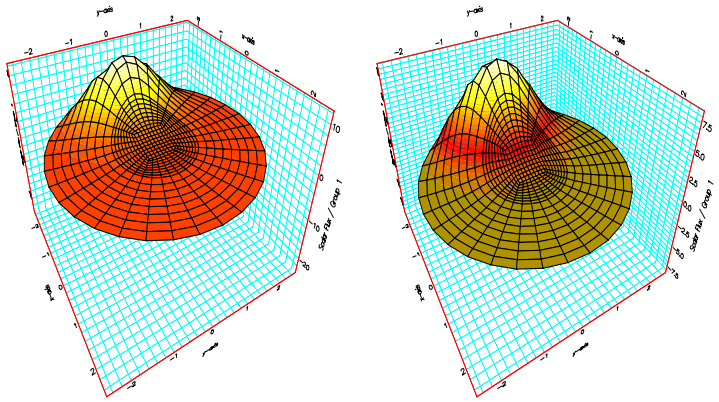
the essential features of the time dependent signal.

Having justified the ability to use large time steps, a P_1 angular approximation may be used to perform image reconstructions using 5 Discontinuous Galerkin time steps. Figure 6.3 depicts scalar flux as a function of two dimensional space (x and y) for the exact solution at time levels 1 (top left hand side graph), 2 (top middle graph), 3 (top right hand side graph), 4 (bottom left hand side graph), and 5 (bottom right hand side graph) respectively. The delta function signal starts off as a large, slim peak before propagating through the medium as time increases. Figure 6.3 shows how the scalar flux disperses through the medium and how the peak scalar flux decreases as a function of time.

Figure 6.4 illustrates scalar flux as a function of x- and y- spatial co-ordinates for the forward solution (left hand side graph), and the adjoint solution (right hand side graph) for a



Time Levels 1 → 2 → 3



Time Levels 4 → 5

Figure 6.3: Scalar flux as a function of two dimensional space (x and y) for the exact solution at time levels 1 (top left hand side graph), 2 (top middle graph), 3 (top right hand side graph), 4 (bottom left hand side graph), and 5 (bottom right hand side graph). A P_3 angular approximation has been used to obtain each of the above images. The scalar flux extends up to a maximum magnitude of $125 cm^{-2}s^{-1}$ during the first time step, $250 cm^{-2}s^{-1}$ during the second time step, $125 cm^{-2}s^{-1}$ during the third time step, $10 cm^{-2}s^{-1}$ during the fourth time step and $7.5 cm^{-2}s^{-1}$ during the fifth time step.

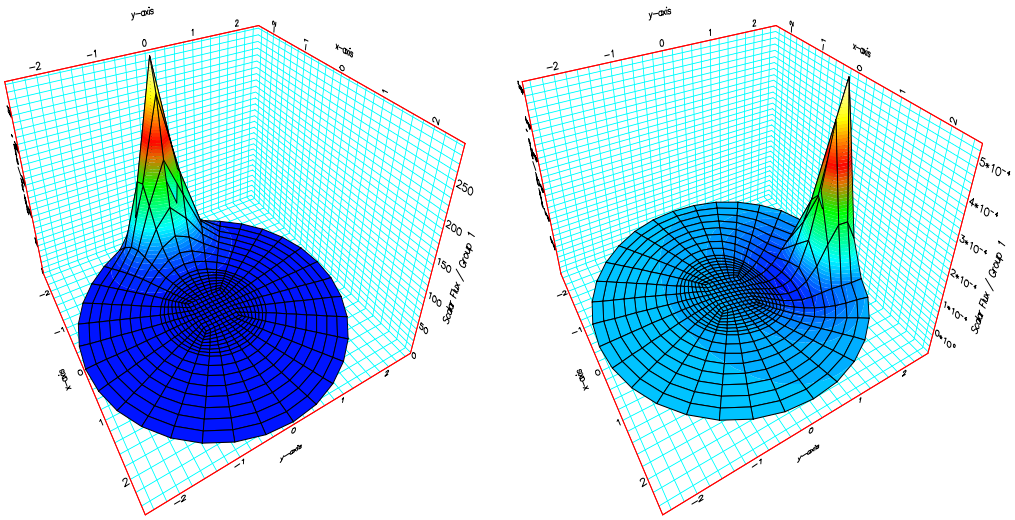


Figure 6.4: Scalar flux as a function of two dimensional space (x and y) for the forward solution (left hand side graph), and the adjoint solution (right hand side graph). A P_1 angular approximation is used to obtain each of the above images.

single source-detector arrangement. The peak illustrated in the forward solution diagram indicates the position of the isotropic source (at co-ordinates -2.5, 0.0). The adjoint solution represents the data misfit between the exact solution and the model prediction. The adjoint solution illustrates the adjoint flux obtained from the third detector placed at co-ordinates (0.0, -2.5). The first and second order gradient of the error functional attempts to adjust the absorption and scattering material properties so that the forward model can better fit the exact solution obtained at the detectors. As a result the scalar flux of the adjoint solution, the error functional, and the gradient of the error functional all tend to zero as a more accurate angular flux approximation is found.

Figure 6.5 compares the absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 6.1 and table 6.2. A P_1 angular approximation was used to perform the transient forward model and inversion model calculations. The inversion was performed for 11 iterations.

Figure 6.5 provides visual evidence that during the P_1 inversion calculation, the algorithm finds the correct location and approximate size of the inhomogeneity for both absorption and scattering reconstructions. A P_1 angular approximation therefore provides an adequate description for modelling time dependent radiation transport in highly scattering

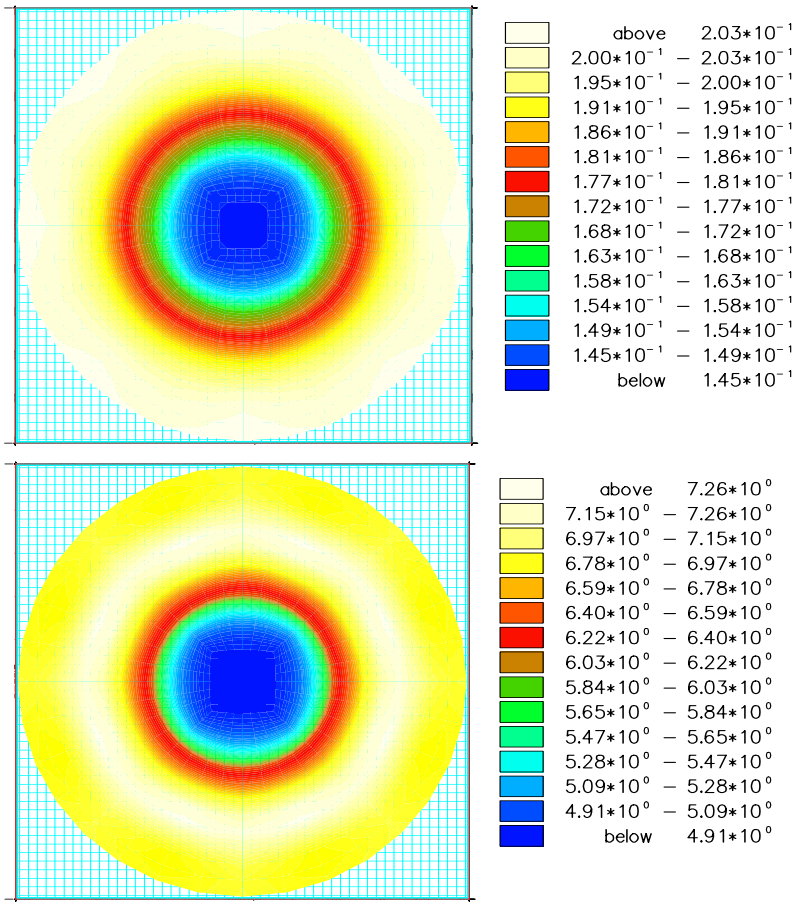


Figure 6.5: Absorption (top image) and scattering (bottom image) reconstruction results for a circle in circle type problem using a P_1 angular expansion. The image illustrates the first 11 iterations of the transient reconstruction.

media. The relative magnitudes of the absorption and scattering coefficients are also approximated fairly accurately. The magnitude of the absorption coefficient, μ_a, extends from 0.2cm^{-1} to approximately 0.145cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 7.26cm^{-1} to approximately 4.91cm^{-1}. The magnitude of the absorption coefficient of the inhomogeneity is slightly overestimated, while the magnitudes of the background medium and the scattering coefficient remain accurate.

Figure 6.6 illustrates the error functional as a function of the iteration number for the P_1 angular approximation described above. The top graph illustrates the reduction of the error functional during the first 11 iterations plotted on a log scale. The bottom graph illustrates the same results plotted on a linear scale. Figure 6.6 illustrates that the error

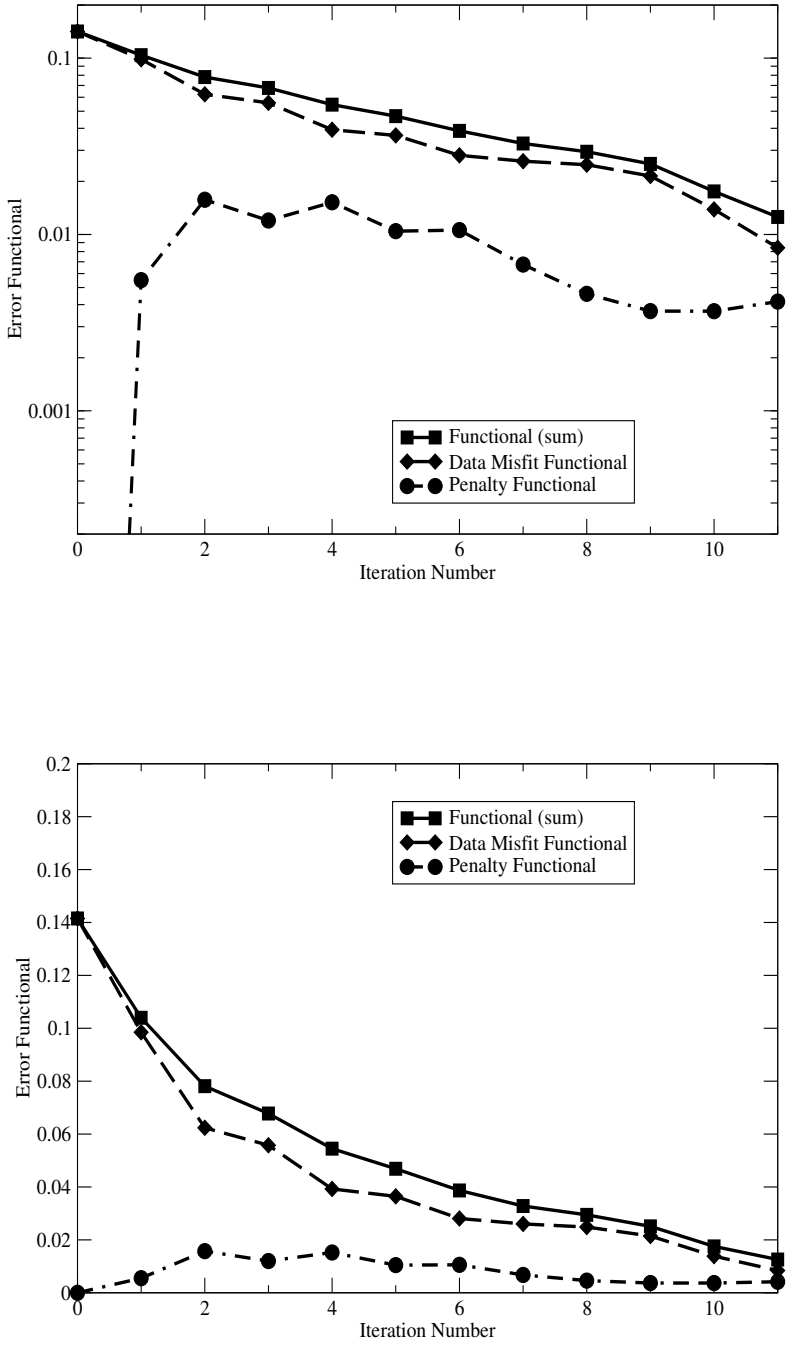
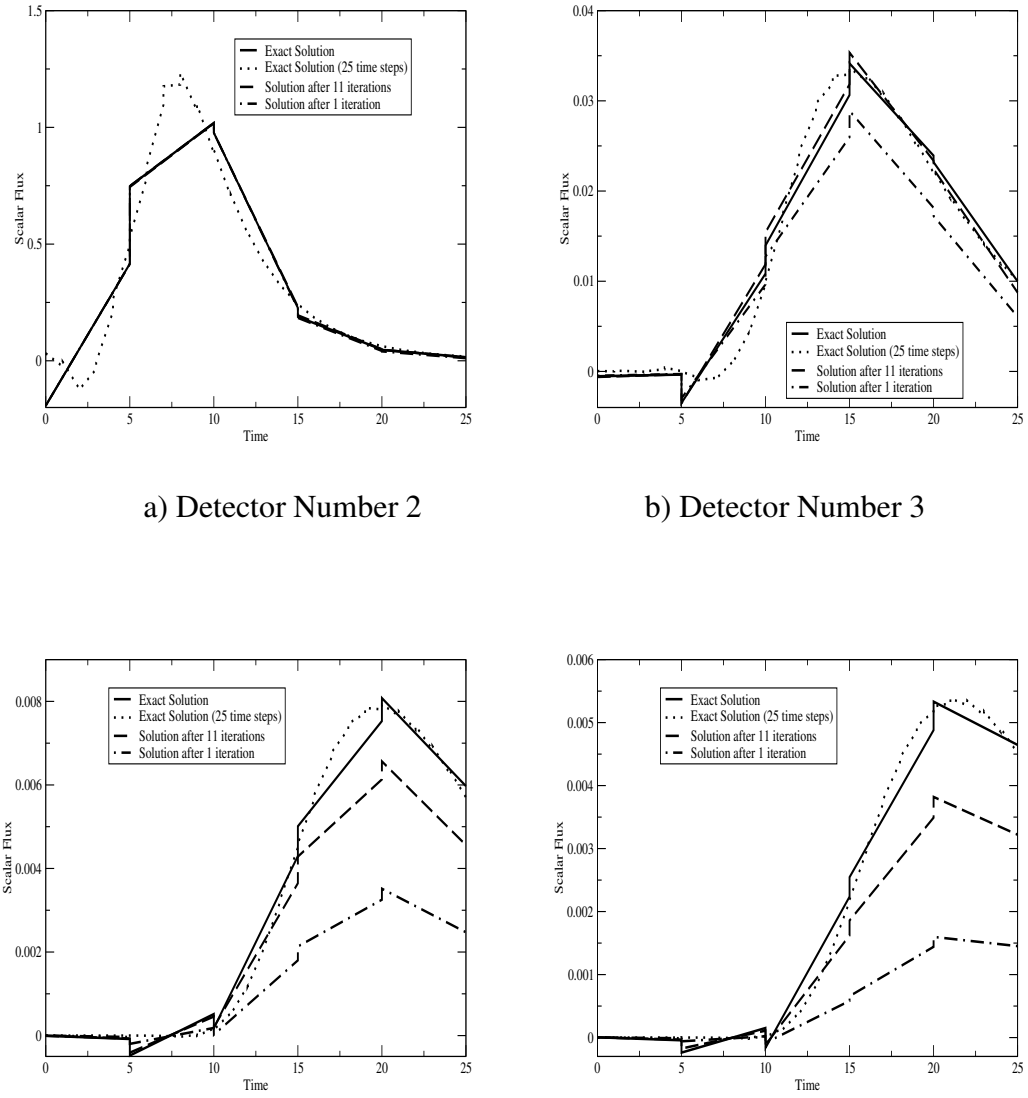


Figure 6.6: Error functional as a function of iteration number. The top graph illustrates the reduction of the error functional during the first 11 iterations plotted on a log scale. The bottom graph illustrates the same results plotted on a linear scale.



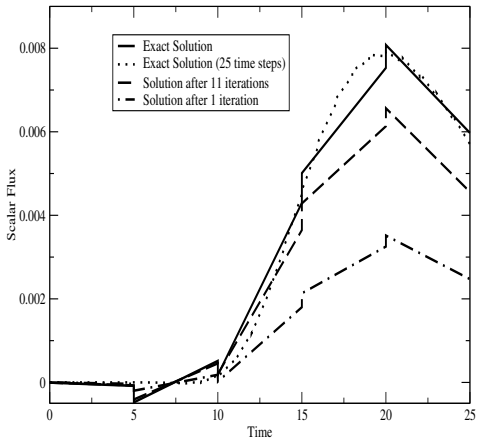
a) Detector Number 2

b) Detector Number 3

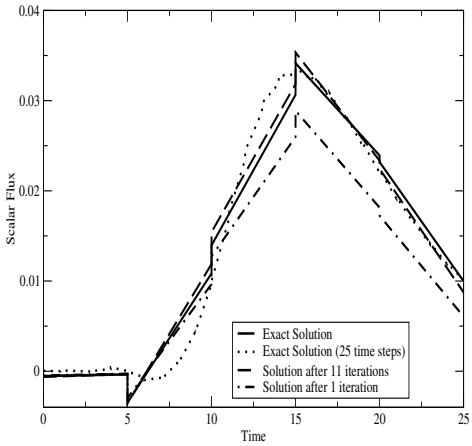
c) Detector Number 4

d) Detector Number 5

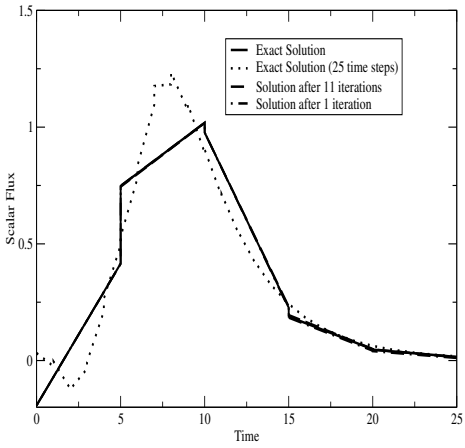
Figure 6.7: Scalar flux versus time for the first source at a) the second detector location, b) the third detector location, c) the fourth detector location and d) the fifth detector location. The graphs compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 11 iterations (dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each graph.



a) Detector Number 6



b) Detector Number 7



c) Detector 8

Figure 6.8: Scalar flux versus time for the first source at a) the sixth detector location, b) the seventh detector location, and c) the eighth detector location. The graphs compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 11 iterations (dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each graph.

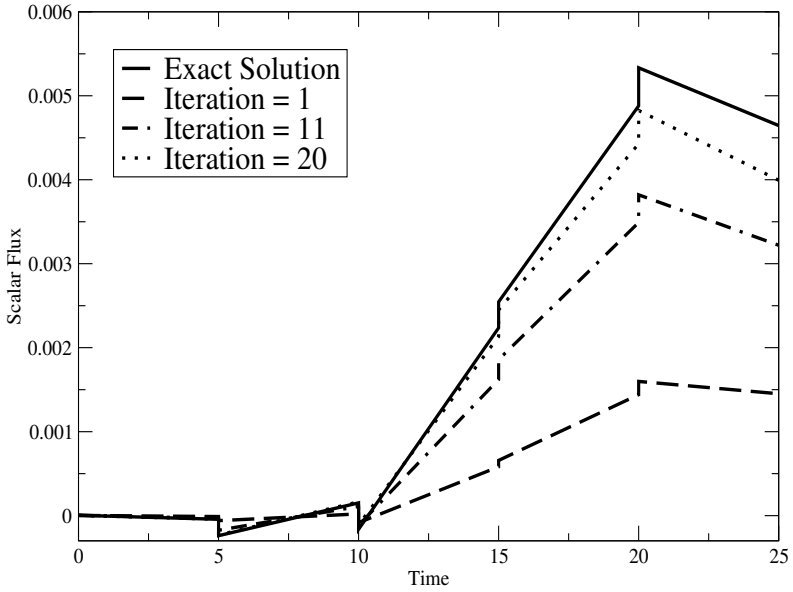


Figure 6.9: Scalar flux versus time for the first source and the fifth detector position. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), the forward model result after 11 iterations (dot-dashed line), and the forward model result after 20 iterations (dotted line). The solution approaches convergence as the number of iterations is increased.

functional is reduced most dramatically during early stages of the inversion (i.e. during the first six iterations). The error functional is reduced by more than an order of magnitude during the first 11 iterations, from a functional of magnitude $F = 0.14$ to a functional of magnitude $F = 0.013$.

Figure 6.7 illustrates scalar flux as a function of time for the first source at the second detector location (top left hand graph), the third detector location (top right hand graph), the fourth detector location (bottom left hand graph) and fifth detector location (bottom right hand graph), while figure 6.8 illustrates scalar flux as a function of time for the first source at the sixth detector location (top left hand graph), the seventh detector location (top right hand graph), and the eighth detector location (bottom left hand graph). The eight isotropic sources and detectors are placed equidistantly around the circumference of the medium.

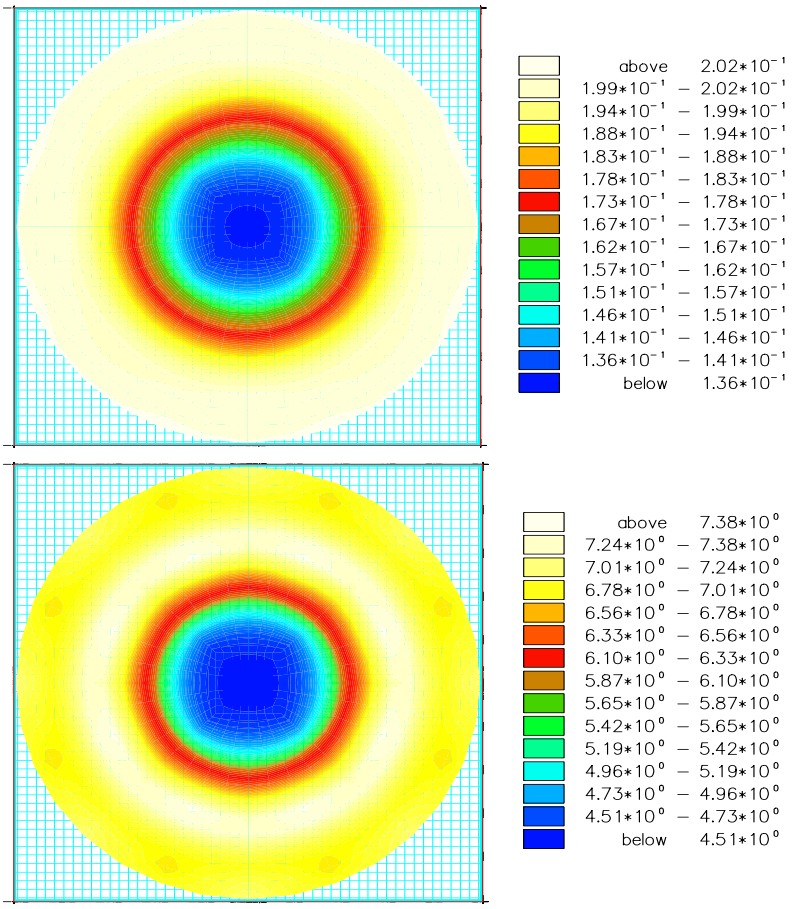


Figure 6.10: Absorption (top image) and scattering (bottom image) reconstruction results for a circle in circle type problem using a P_1 angular expansion. The figure illustrates the first 20 iterations of the time-dependent reconstruction scheme.

The co-ordinate positions of each of the individual sources and detectors are described in table 6.1. Figures 6.7 and 6.8 compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 11 iterations (dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in figures 6.7 and 6.8. The detector response received at the first detector has been omitted from the calculations (using the weight function) because the large source-detector sensitivities observed when sources and detectors are placed too close together are not regarded as reliable. The first detector has been placed on top of the first source.

Important features illustrated in figures 6.7 and 6.8 include the symmetry of the modelling

domain and the convergence characteristics of the inverse problem. The similarity in scalar flux versus time graphs at the second and eighth detector locations, the third and seventh detector locations and the fourth and sixth detector locations indicate that the model is symmetric and that the sources and detectors have been placed equidistantly around the circumference of the medium. In other words, the second and eighth detectors, third and seventh detectors and the fourth and sixth detectors are placed at equal distances away from the source. Thus the signal received at each of the detector pairs remains the same.

The convergence characteristics show that as the number of iterations is increased the detector response at each of the source-detector pairs approaches that of the exact solution. However, the graphs illustrated in figures 6.7 and 6.8 suggest that the model has some way to go before convergence is met. This is illustrated most dramatically in the detector response observed at the fifth detector. The fifth detector is placed furthest away from the source and therefore has the weakest detector response. Therefore, because of the absence of useful data, the scalar flux versus time graph observed at the fifth detector remains the most difficult signal to fit to an exact solution.

Convergence is improved when the number of iterations is increased. Figure 6.9 depicts scalar flux as a function of time at the first source and the fifth detector positions. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), the forward model result after 11 iterations (dot-dashed line), and the forward model result after 20 iterations (dotted line). A marked improvement in the convergence of the scalar flux is illustrated as the number of iterations is increased to 20. After 20 iterations the error functional is reduced to a magnitude of $F = 0.0023$ and convergence is reached. The convergence criterion is chosen so that once convergence is met, increasing the number of iterations does not change the inversion model significantly.

Figure 6.10 compares the absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 6.1 and table 6.2. A P_1 angular approximation was used to perform the transient forward model and inversion model calculations and convergence was met after 20 Levenberg-Marquardt iterations. The relative magnitude of the absorption coefficient shows a slight improvement when

compared to Levenberg-Marquardt calculations performed for 11 iterations. The magnitude of the absorption coefficient, μ_a, extends from 0.2cm^{-1} to approximately 0.138cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 7.26cm^{-1} to approximately 4.51cm^{-1}. The relative magnitude of the reconstructed absorption coefficient still remains slightly overestimated.

The reasoning behind initially using 11 iterations to perform the time dependent image reconstruction is that it took 11 Levenberg-Marquardt iterations to reach convergence for steady state inversion calculations using the same problem set up and model geometry (see [43]). Although imaging using time dependent information may take a larger number of iterations to converge, the image reconstruction results produce sharper images (see [43] for comparison) and the relative absorption and scattering coefficient magnitudes are reconstructed more accurately. Time dependent information can therefore be used to provide better quality inversion images.

6.5.2 Experiment 2

Figure 6.11 illustrates a square domain of size 5cm $\times$ 5cm. The domain contains three embedded inhomogeneities, each of size 1cm $\times$ 1cm . The inhomogeneities are placed 1cm deep inside the medium. The absorption and scattering material properties of both the background medium and the embedded inhomogeneities are described in table 6.3. Two of the embedded inhomogeneities have the same absorption material properties and two of the embedded inhomogeneities have the same scattering material properties. Sixteen equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward and reconstruction model results. A P_1 angular approximation is used to obtain the time-dependent forward and inversion model calculations along with a Levenberg-Marquardt optimisation method. The same sized mesh is used to generate both the two dimensional forward model and inversion model reconstruction. The mesh has 882 nodes, 1280 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100,000 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially

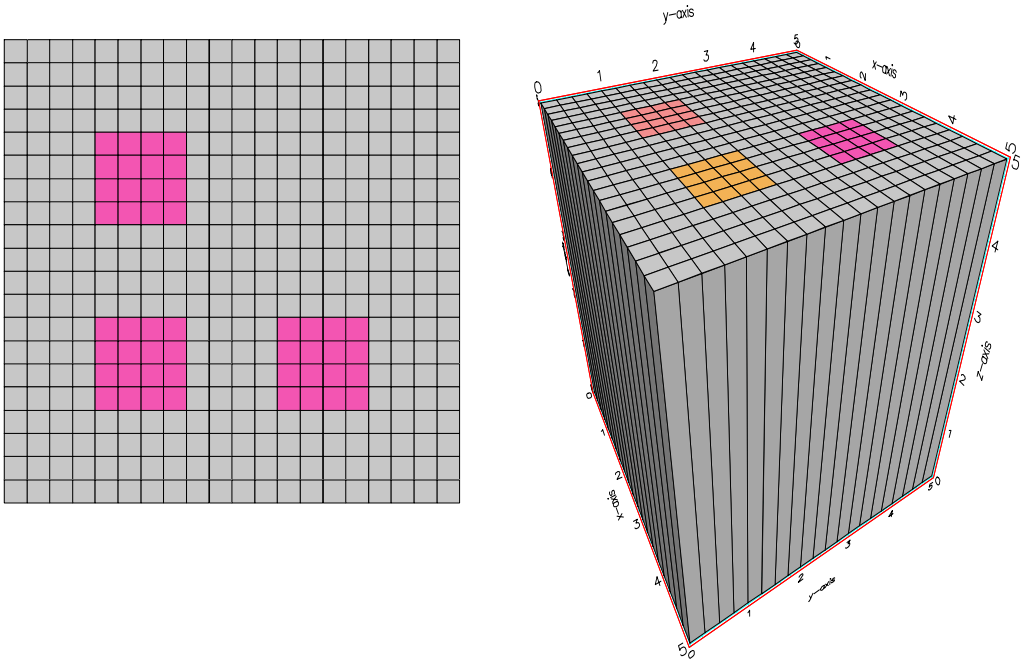


Figure 6.11: A two dimensional structured finite element square mesh for a phantom of size 5cm $\times$ 5cm. Three inhomogeneities are located 1cm deep inside the medium. The mesh used for both the forward model and inversion model reconstruction has 882 nodes, 1280 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps.

set to 0.1. As an initial guess, a homogeneous medium is chosen with optical properties that equal the background. In other words $\mu_a = 0.07$cm and $\mu_s^{'} = 10.0$cm.

Figure 6.12 compares the absorption (top image) and scattering (bottom image) reconstruction results for the (time dependent) problem geometry described in figure 6.11 and table 6.3. A P_1 angular approximation provided an adequate description for modelling radiation transport. Figure 6.12 provides visual evidence that the inversion algorithm finds the correct location, size and approximate shape of the three inhomogeneities for both absorption and scattering reconstructions. The approximate magnitudes of the absorption coefficients and scattering coefficients of each of the three inhomogeneities are also reconstructed relatively accurately. The P_1 results approximate the magnitude of the absorption coefficient, μ_a, to extend from 0.106cm^{-1} to approximately 0.067cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 16.4cm^{-1} to approximately 8.93cm^{-1}. The magnitude of the background absorption and scattering coefficients remain accurate.

	$\mu_a[cm^{-1}]$	$\mu_s'[cm^{-1}]$
Background Medium	0.07	10.0
Top Left Inhomogeneity	0.07	15.0
Bottom Left Inhomogeneity	0.14	10.0
Bottom Right Inhomogeneity	0.14	15.0

Table 6.3: Optical absorption and scattering coefficients for a square domain of size 5cm $\times$ 5cm. Three inhomogeneities are contained inside the domain. The scattering and absorption material properties of each of the inhomogeneities are described.

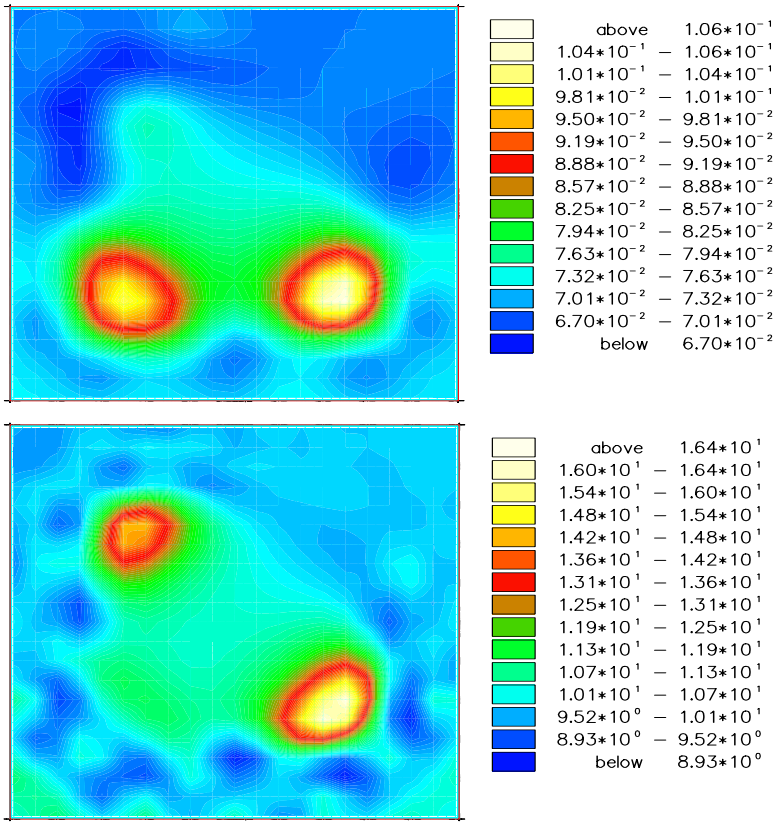


Figure 6.12: Absorption (top image) and scattering (bottom image) coefficient reconstruction for a square domain problem containing three embedded inhomogeneities. Two of the embedded inhomogeneities have the same absorption material properties and two of the embedded inhomogeneities have the same scattering material properties. The transient inversion was performed using a P_1 angular approximation and was run for 43 iterations.

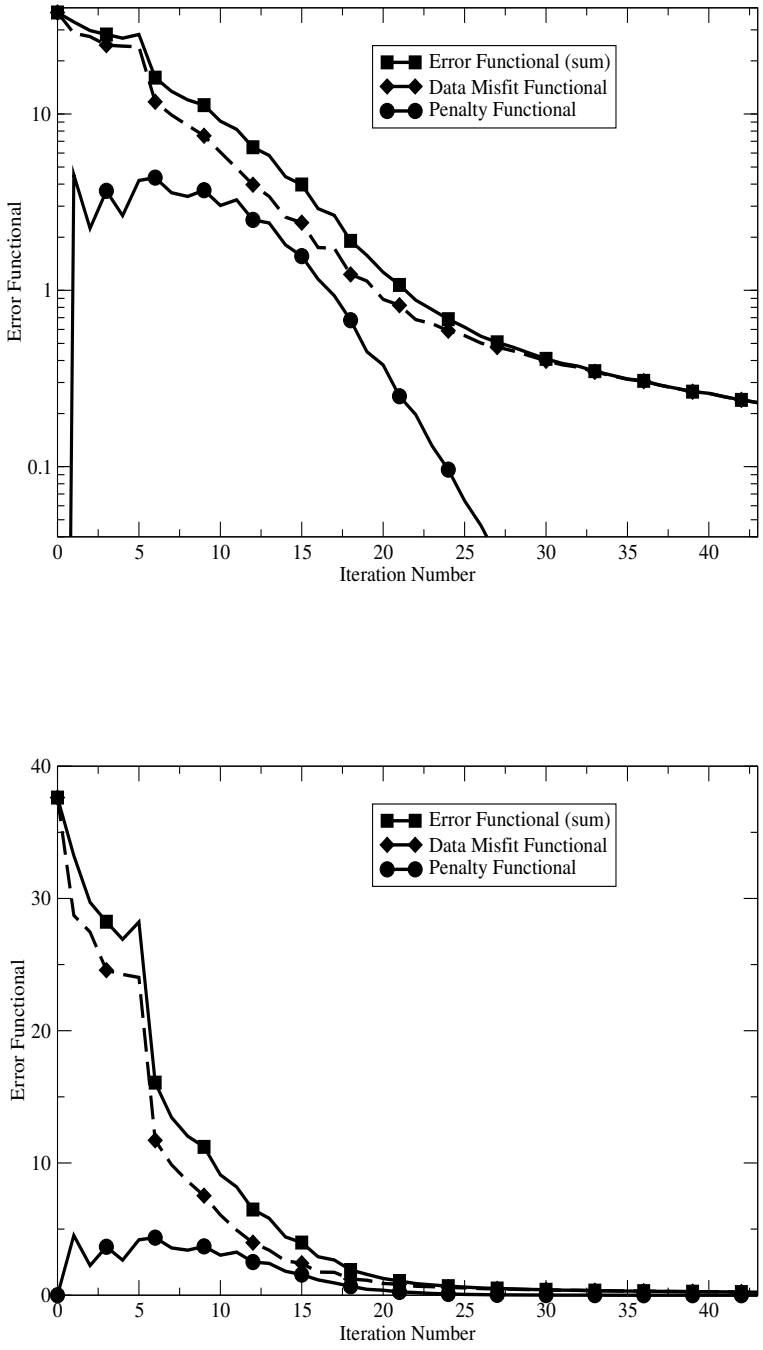


Figure 6.13: Error functional as a function of iteration number for time dependent radiation transport. The top graph illustrates the reduction of the error functional during the first 43 iterations plotted on a log scale. The bottom graph illustrates the same results plotted on a linear scale. The gradual decline in the error functional is because γ_a and γ_{s0} are initially set to large magnitudes.

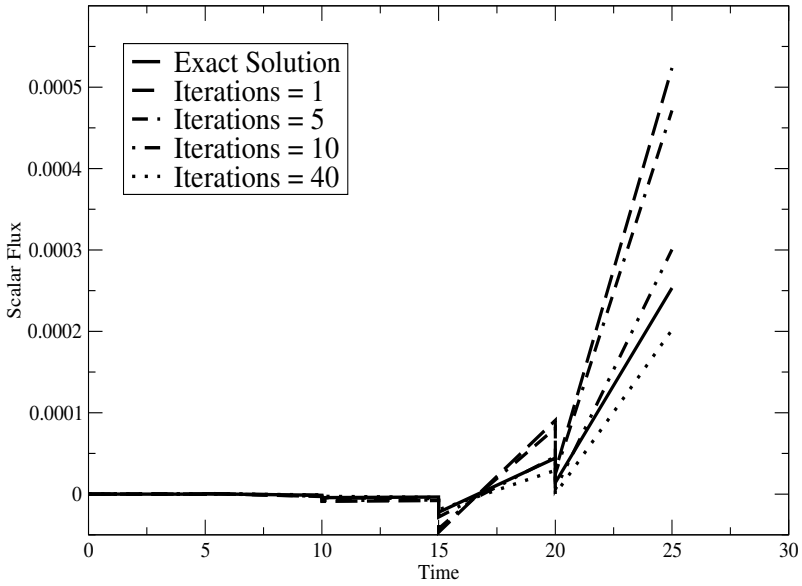


Figure 6.14: Scalar flux versus time for the first source and the eighth detector position. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), the forward model result after 5 iterations (dot-dashed line), the forward model result after 10 iterations (dash with two dots) and the forward model results after 40 iterations (dotted line). The solution approaches convergence as the number of iterations is increased.

Like image reconstruction using steady state radiation transport theory (chapter 5), the algorithm manages to distinguish between absorption and scattering features with relatively little cross-talk. However, the inversion scheme that utilises time dependent information resolves the approximate sizes and shapes of the three embedded inhomogeneities more accurately and produces a sharper image reconstruction for both scattering and absorption material properties. The relative absorption and scattering coefficient magnitudes are also reconstructed more accurately.

Figure 6.13 illustrates error functional as a function of iteration number for the first 43 iterations of the P_1 angular expansion. The model reduced the error functional from a magnitude of $F = 37.626$ to a magnitude of $F = 0.231$. The regularisation penalty levels

are initially set to large initial values of $\gamma_a = \gamma_{s0} = 100,000$ and are annealed downwards by a factor of 1.5 after each iteration. The small relative reduction in γ_a and γ_{s0} provides a gradual decrease in the error functional.

The converging solution is illustrated in figure 6.14. Figure 6.14 depicts scalar flux as a function of time at the first source and the eighth detector positions. The x- and y- position co-ordinates of each individual source and detector are illustrated in table 6.1. The first source is placed at coordinates (1.0, 0.0) and the eighth detector is placed furthest away from the source at coordinates (5.0, 4.0). The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), the forward model result after 5 iterations (dot-dashed line), the forward model result after 10 iterations (dashed line with two dots) and the forward model result after 40 iterations (dotted line). A marked improvement in the convergence of the scalar flux is illustrated as the number of iterations is increased.

6.5.3 Experiment 3

The simulation geometry described in figure 6.15 depicts a circle 7cm in diameter containing an embedded void-like ring of thickness 0.5cm. The ring extends from a radius of 2cm to a radius of 2.5cm from the centre of the domain. The absorption and scattering coefficients of the background medium and the inhomogeneous ring are contained in table 6.4. Eight equally spaced isotropic sources and detectors are placed around the circumference of the medium to obtain the forward model and inverse model reconstruction results. A P_1 angular approximation is used to obtain the time dependent forward and inversion calculations along with a Levenberg-Marquardt optimisation method. The same sized mesh is used to generate both the two dimensional forward model and inversion model reconstruction. The mesh used has 2210 nodes, 3296 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 10,000 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium is chosen with optical properties that equal the background. In other words $\mu_a = 0.2$cm and $\mu_s^{'} = 7.5$cm.

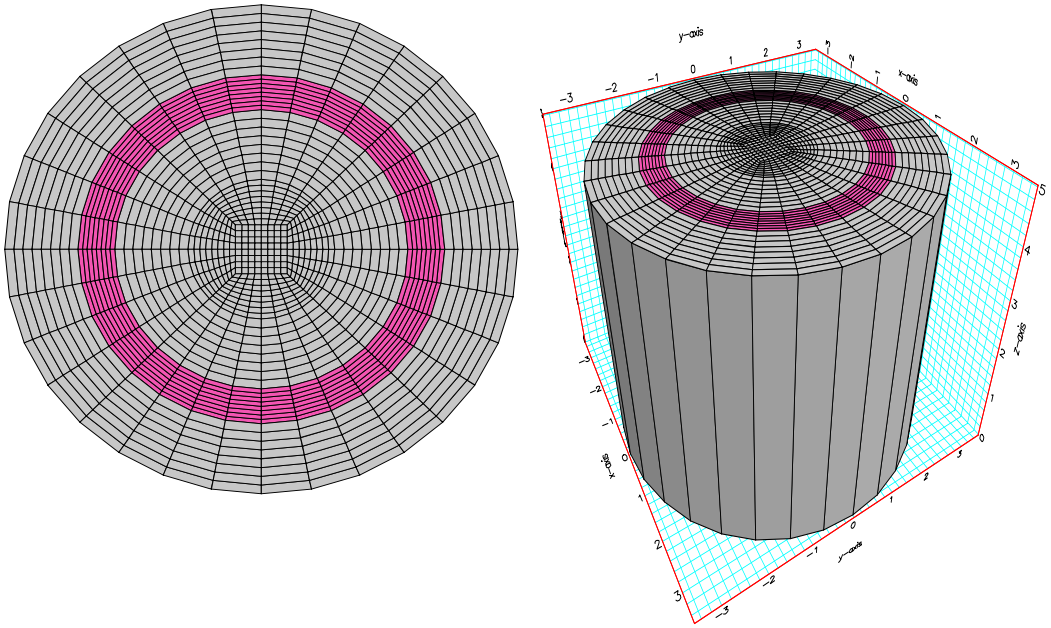


Figure 6.15: A two dimensional structured finite element circular mesh of diameter 7cm with an embedded void ring of thickness 0.5cm. The ring extends from a radius of 2cm to a radius of 2.5cm from the centre of the domain. The time dependent problem has 2210 nodes, 3296 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.2	7.5
Ring Region	0.0	0.0

Table 6.4: Optical absorption and scattering coefficients for a circle 7cm in diameter containing an embedded void-like ring of thickness 0.5cm.

Figure 6.16 compares the converged absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 6.15 and table 6.4. A P_1 angular approximation was used to perform the transient forward model and inversion model calculations. The inversion was performed for 42 iterations. Figure 6.16 provides visual evidence that during the P_1 inversion calculation, the algorithm finds the correct location and approximate size of the transparent ring for both absorption and scattering reconstructions.

The relative magnitudes of the absorption and scattering coefficients are also approximated fairly accurately although are never absolutely resolved. The magnitude of the absorption coefficient, μ_a, extends from 0.183cm^{-1} to approximately 0.029cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 15.3cm^{-1} to approximately 0.115cm^{-1}. The magnitudes of both the absorption and scattering coefficient reconstructions indicate that a void is present but are both overestimated inside the transparent void. The magnitude of the background scattering coefficient is also overestimated in regions where eight image artefacts exist. The eight image artefacts are located in the vicinity of where the eight sources and eight detectors are situated. The scattering coefficient magnitudes of these regions could be improved either by increasing the regularisation penalty levels, γ_a and γ_{s0}, or by utilising a more sophisticated regularisation strategy.

A Large improvement in image quality is observed when comparing calculations performed using time dependent information to calculations performed using steady state information. Figure 6.17 compares the converged absorption (left hand side) and scattering (right hand side) reconstruction results for the problem geometry described in figure 6.15 and table 6.4 using steady state radiation transport. P_1 (top) and P_3 (bottom) angular approximations were used to perform the inversion calculations. Figure 6.17 provides visual evidence that during the steady state P_3 scattering reconstruction, the algorithm finds the correct location of the transparent ring. However, the thickness of the void is overestimated. The P_3 angular approximation fails to reconstruct the absorption material structure. The P_1 angular approximation fails to reconstruct both the scattering and absorption material structure. The results suggest that steady state radiation transport theory has limited use when considering void modelling. Utilising time dependent transport

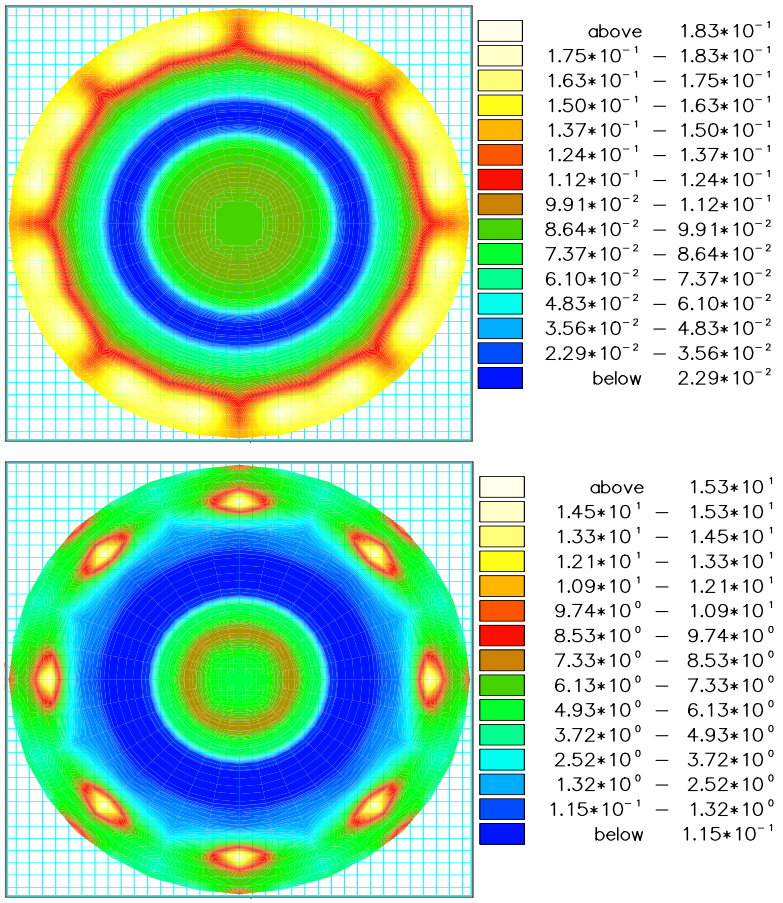


Figure 6.16: Absorption (top image) and scattering (bottom image) reconstruction results for a circular problem containing an embedded transparent ring. The figure illustrates a time-dependent reconstruction using a P_1 angular expansion. The inversion converged after 42 iterations.

could provide the inversion model with crucial transient information.

The relative magnitudes of the steady state absorption and scattering coefficients are approximated fairly accurately although they are never absolutely resolved. The P_1 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.182cm^{-1} to approximately 0.0147cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.41cm^{-1} to approximately 0.0871cm^{-1}. The P_3 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.214cm^{-1} to approximately 0.0232cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.41cm^{-1} to approximately 0.527cm^{-1}. The legend

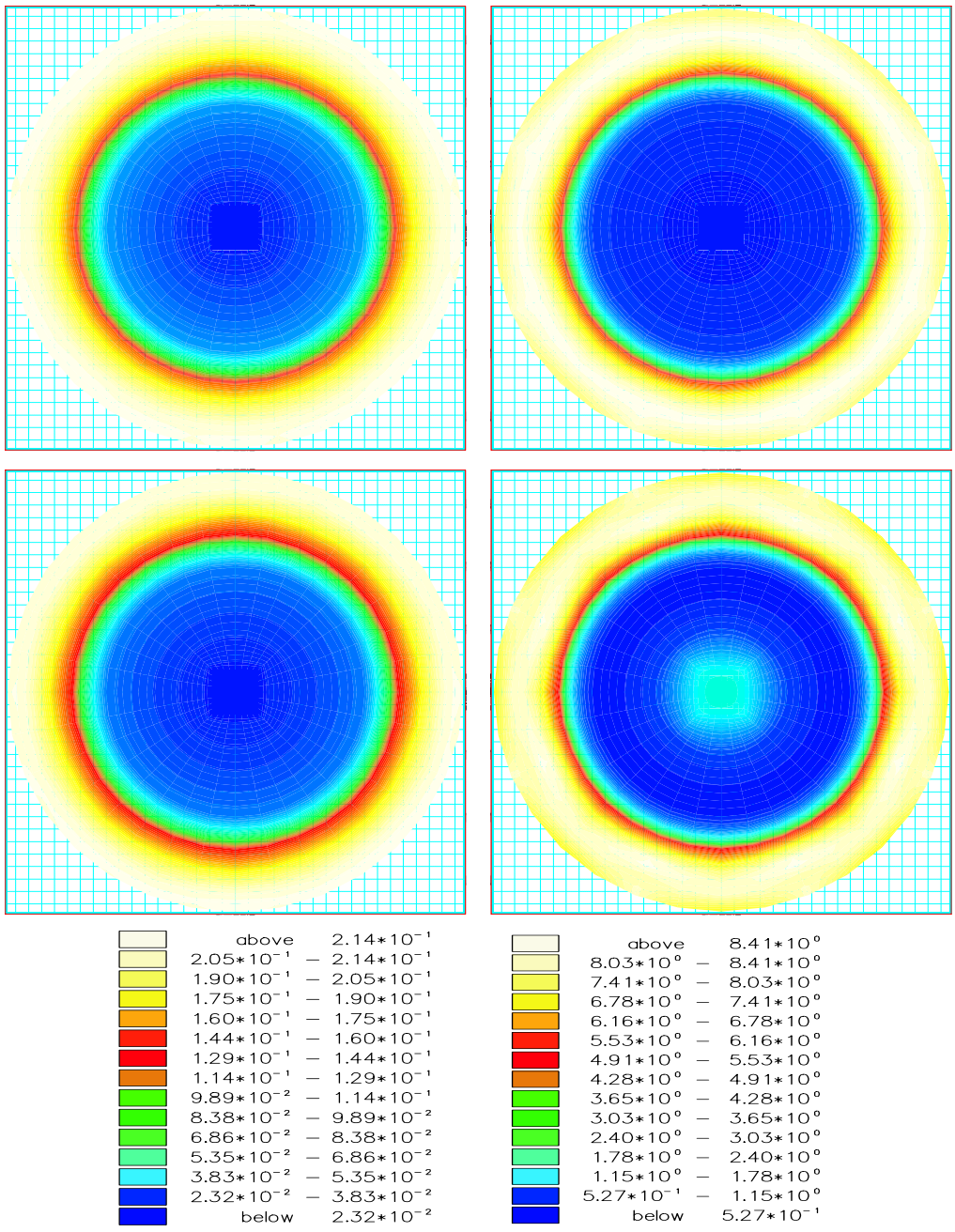


Figure 6.17: Steady state absorption (left hand side) and scattering (right hand side) reconstruction results for a circular problem containing an embedded void ring. P_1 (top) and P_3 (bottom) angular approximations are used to reconstruct the medium's material properties. The P_1 angular approximation was performed for 13 iterations and the P_3 angular approximation was performed for 24 iterations. The legend depicts the range in magnitude of the absorption coefficient (left hand side) and scattering coefficient (right hand side) for the P_3 angular approximation.

illustrated at the bottom of figure 6.17 depicts the range in magnitude of the absorption coefficient (left hand side) and scattering coefficient (right hand side) for the P_3 angular approximation described above. For both P_1 and P_3 angular approximations, the magnitudes of both the scattering and absorption coefficients of the inhomogeneous ring indicate that a void is present but are overestimated. The magnitude of the background scattering coefficient is also overestimated for both P_1 and P_3 angular approximations. The magnitude of the scattering coefficient reconstruction of the internal circle (contained on the inside of the inhomogeneous ring), described in the P_3 reconstruction is slightly underestimated.

Figure 6.18 illustrates the error functional as a function of the iteration number for the time dependent, P_1 angular approximation described above. The top graph illustrates the reduction of the error functional during the first 42 iterations plotted on a log scale. The bottom graph illustrates the same results plotted on a linear scale. The error functional is reduced by two orders of magnitude during the first 42 iterations, from a functional of magnitude $F = 11.60$ to a functional of magnitude $F = 0.545$. Figure 6.18 illustrates a fairly gradual decrease in the error functional. This is because γ_a and γ_{s0} are initially set to large magnitudes and are annealed down gradually.

Figure 6.19 illustrates scalar flux as a function of time for the first source at the second detector location (top left hand graph), the third detector location (top right hand graph), the fourth detector location (bottom left hand graph) and fifth detector location (bottom right hand graph). Figure 6.20 illustrates scalar flux as a function of time for the first source at the sixth detector location (top left hand graph), the seventh detector location (top right hand graph), and the eighth detector location (bottom left hand graph). The eight isotropic sources and detectors are placed equidistantly around the circumference of the domain. The co-ordinate positions of each of the individual sources and detectors are described in table 6.1. The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in figures 6.19 and 6.20. Figures 6.19 and 6.20 compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 40 iterations (dashed line). The detector response received at the first detector has been omitted from the calculations (using the weight function) because the large source-detector sensitivities observed when sources

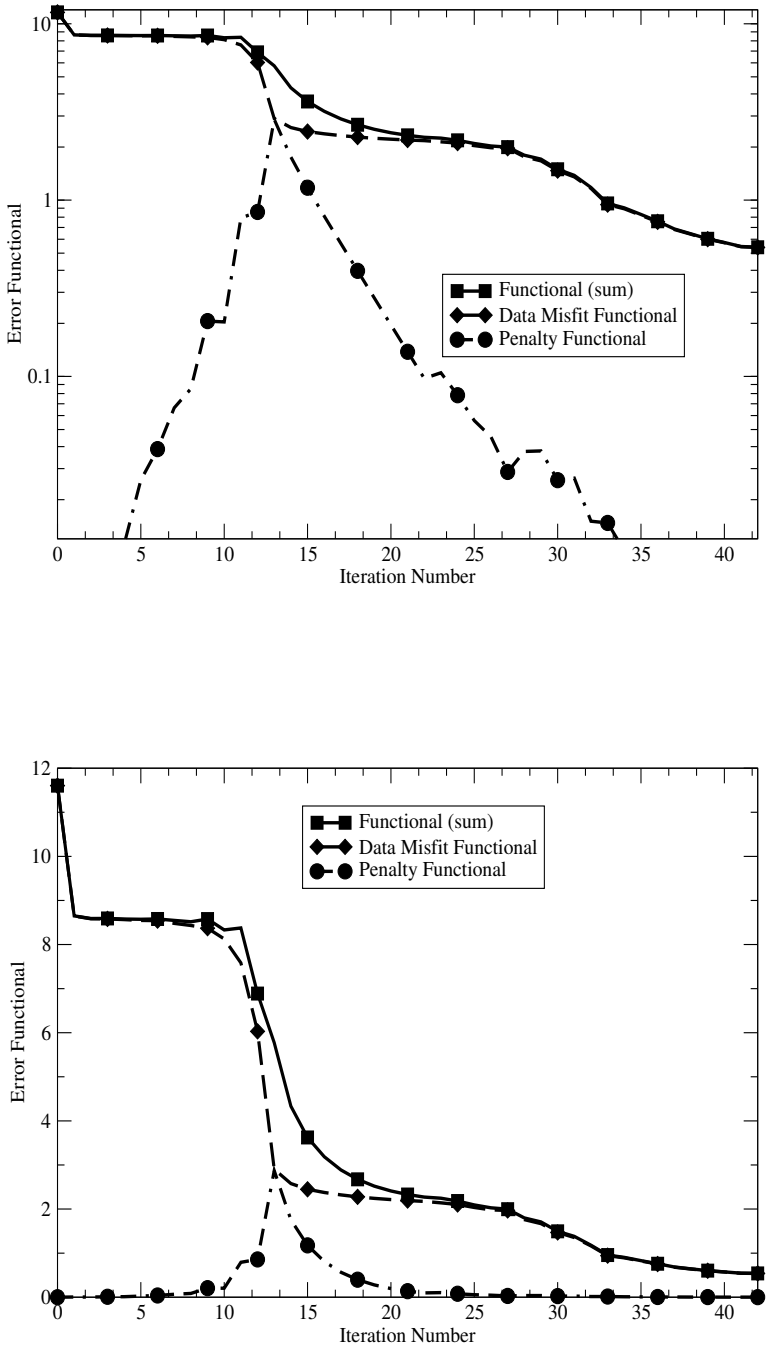
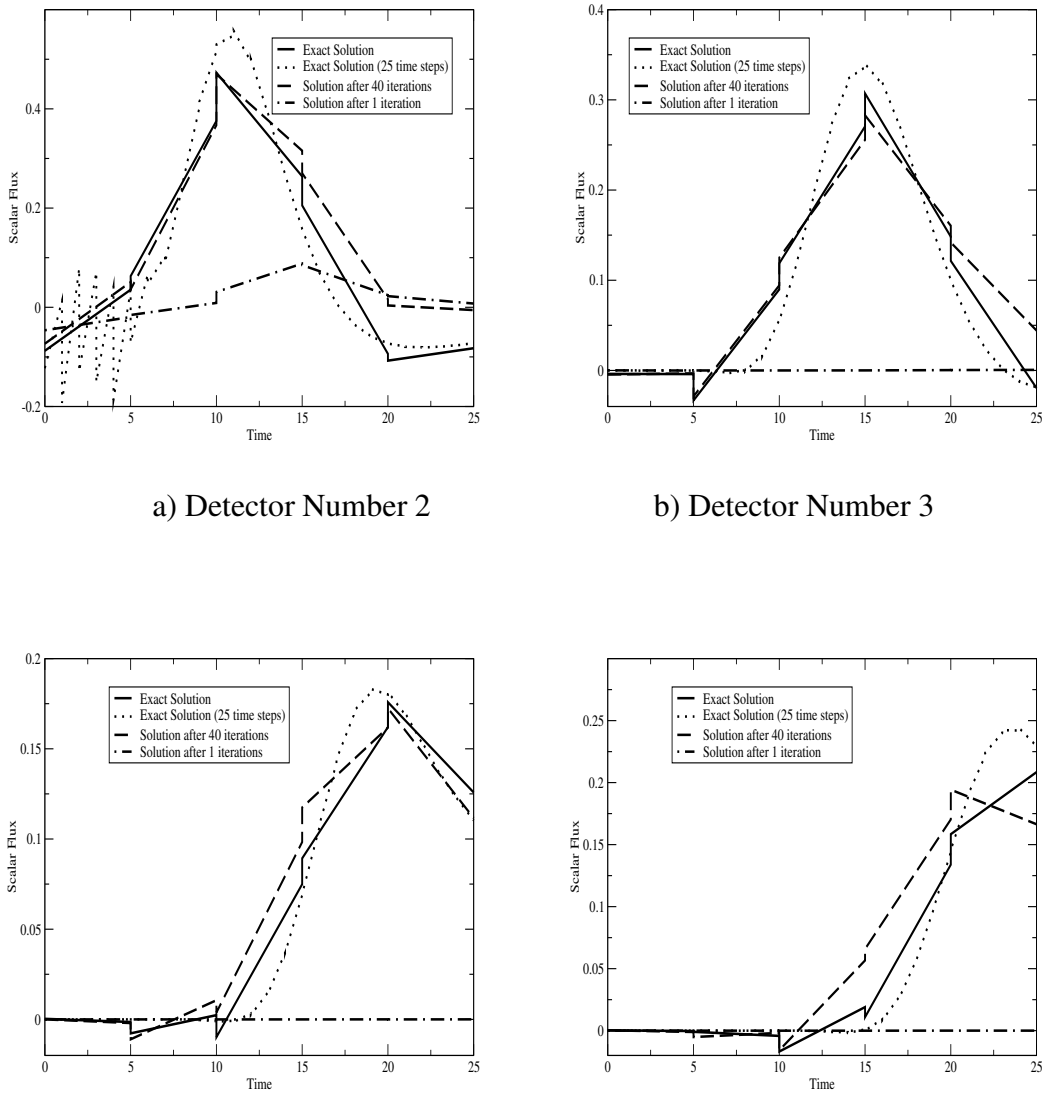


Figure 6.18: Error functional as a function of iteration number for time dependent radiation transport. The top graph illustrates the reduction of the error functional during the first 42 iterations plotted on a log scale. The bottom graph illustrates the same results plotted on a linear scale. The gradual decline in the error functional is because γ_a and γ_{s0} are initially set to large magnitudes.



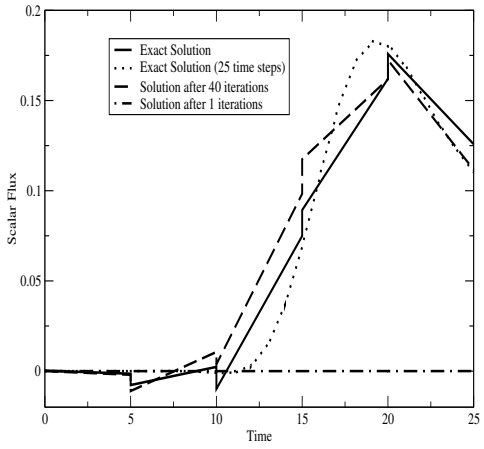
a) Detector Number 2

b) Detector Number 3

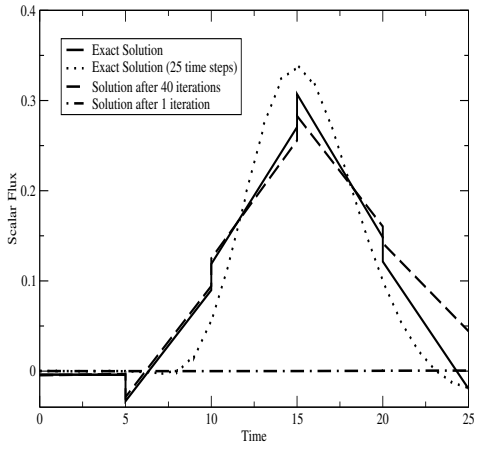
c) Detector Number 4

d) Detector Number 5

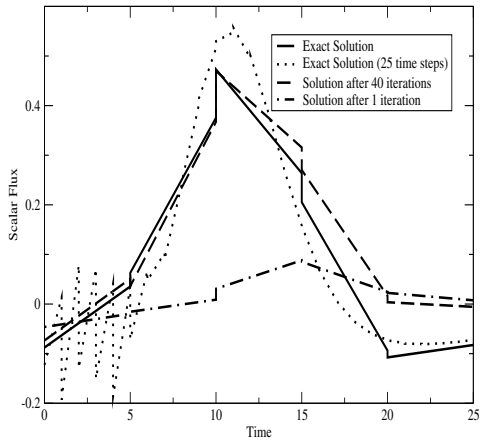
Figure 6.19: Scalar flux versus time for the first source at a) the second detector location, b) the third detector location, c) the fourth detector location and d) the fifth detector location. The graphs compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 40 iterations (dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each graph.



a) Detector Number 6



b) Detector Number 7



c) Detector Number 8

Figure 6.20: Scalar flux versus time for the first source at a) the sixth detector location, b) the seventh detector location, and c) the eighth detector location. The graphs compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 40 iterations (dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each graph.

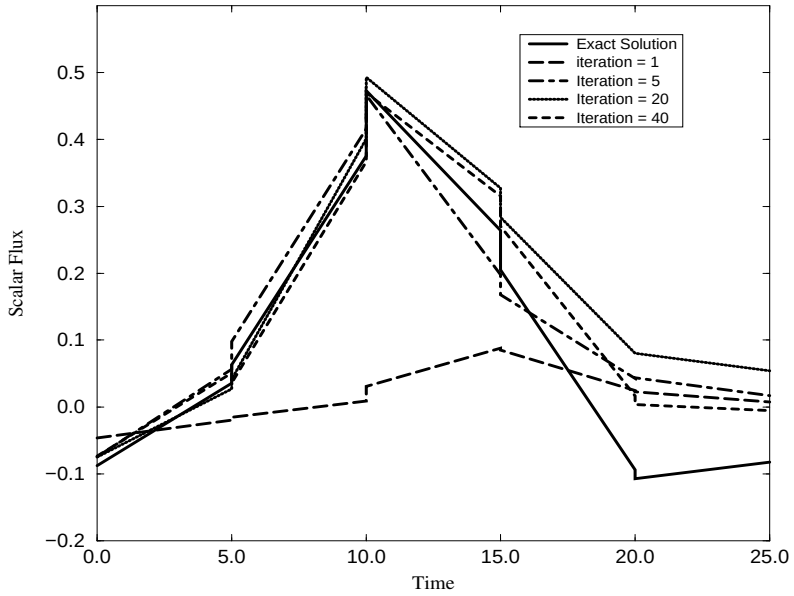


Figure 6.21: Scalar flux versus time for the first source and the second detector position. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), the forward model result after 5 iterations (dot-dashed line), the forward model result after 20 iterations (dotted line) and the forward model results after 40 iterations (dashed line). The solution approaches convergence as the number of iterations is increased.

and detectors are placed too close together are not regarded as reliable. The first detector has been placed on top of the first source.

Important features illustrated in figures 6.19 and 6.20 include the symmetry of the modelling domain and the convergence characteristics of the inverse problem. The similarity in scalar flux versus time graphs at the second and eighth detector locations, the third and seventh detector locations and the fourth and sixth detector locations indicate that the model is symmetric and that the sources and detectors have been placed equidistantly around the circumference of the medium. In other words, the second and eighth detectors, third and seventh detectors and the fourth and sixth detectors are placed at equal distances away from the source. Thus the signal received at each of the detector pairs remains

the same. Another interesting feature observed during the second and eighth detector response is the peak in the scalar flux during the initial part of the time sequence. This peak is observed when calculations are performed for 25 Discontinuous Galerkin time steps and represents the short amount of time the source pulse is activated and the integration of the spatial resolution. In other words the peak is a numerical representation of the delta function in time. The source has been placed on the first of five Discontinuous Galerkin time units.

The convergence characteristics show that as the number of iterations is increased the detector response at each of the source-detector pairs approaches that of the exact solution. After 40 iterations the scalar flux graphs at each of the source-detector pairs converges to that of the exact solution. The convergence criterion is chosen so that once convergence is met, increasing the number of iterations does not change the inversion model significantly. Figure 6.21 depicts scalar flux as a function of time at the first source and the second detector positions. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), the forward model result after 5 iterations (dot-dashed line), the forward model result after 20 iterations (dotted line) and the forward model results after 40 iterations (dashed line). A marked improvement in the convergence of the scalar flux is illustrated as the number of iterations is increased to 40.

6.5.4 Experiment 4

The final simulation geometry, described in figure 6.22, is a square domain containing an embedded transparent, void-like ring. The ring has a 1.0cm thickness and is placed 1.0cm deep into the domain. The absorption and scattering material properties of both the background medium and the transparent void are described in table 6.5. Sixteen equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward and reconstruction model results. Both P_1 and P_3 angular approximations are used to obtain the time dependent forward model and inversion model reconstruction using a Levenberg-Marquardt optimisation method. The same sized mesh is used to generate both the two dimensional forward model and inversion model reconstruction. The mesh has 882 nodes, 1280 linear quadrilateral elements and 5 Dis-

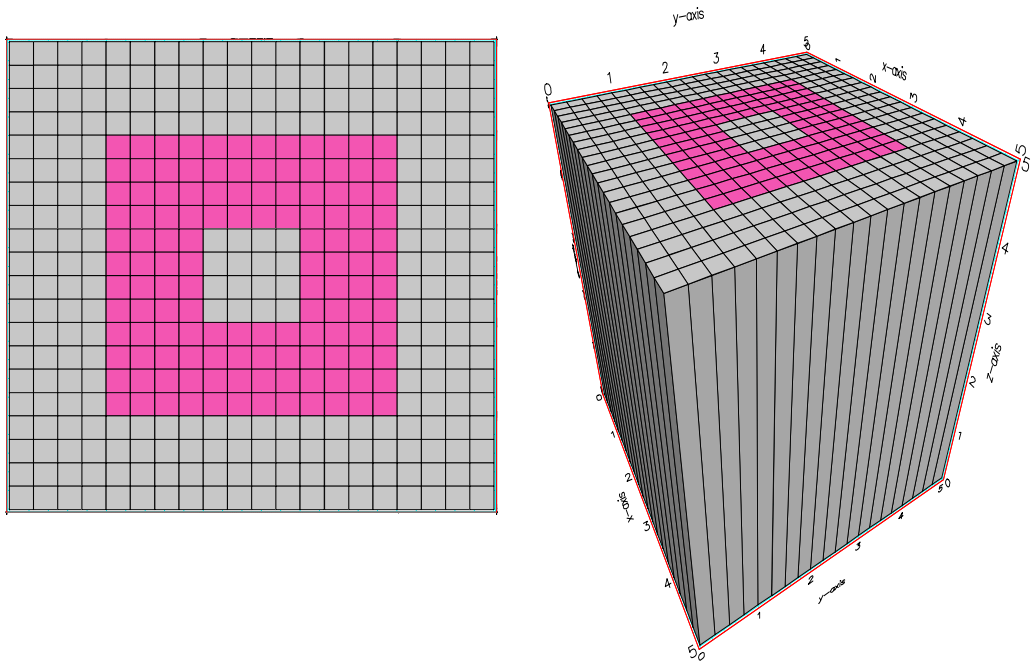


Figure 6.22: A two dimensional structured finite element square mesh for a phantom of size 5cm $\times$ 5cm. A void-like ring of thickness 1.0cm is located at a depth of 1.0cm into the domain. The mesh used for both the forward model and inversion model reconstruction has 882 nodes, 1280 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps.

continuous Galerkin time steps. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium was chosen with optical properties equal to those of the background medium. In other words $\mu_a = 0.07$cm and $\mu_s^{'} = 10.0$cm.

Figures 6.23 and 6.24 compare the absorption (top set of images) and scattering (bottom set of images) reconstruction results, for the problem geometry described in figure 6.22 and table 6.5, after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively. Figure 6.23 illustrates a P_3 angular approximation result, while figure 6.24 illustrates a P_1 angular approximation result. Figure 6.23 provides visual evidence the during the time dependent P_3 scattering reconstruction, the algorithm finds the correct location and thickness of the transparent ring to a high order of accuracy. The P_3 an-

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.07	10.0
Ring Region	0.0	0.0

Table 6.5: Optical absorption and scattering coefficients for a square domain of size 5cm $\times$ 5cm. The square domain contains a transparent void region in the shape of a ring.

gular approximation also goes some way to resolving the ring-like structure during absorption coefficient reconstructions. The P_1 angular approximation fails to reconstruct the ring-like structure for both time dependent scattering and absorption coefficient reconstructions. According to the P_1 images described in figure 6.24, the ring-like structure described in the original problem geometry (figure 6.22) does not exist and instead a large square void is reconstructed.

The relative magnitudes of the absorption and scattering coefficients are approximated fairly accurately although they are never absolutely resolved. The P_1 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.11cm^{-1} to approximately 0.0147cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.8cm^{-1} to approximately 0.166cm^{-1}. The P_3 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.094cm^{-1} to approximately 0.0173cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.98cm^{-1} to approximately 0.18cm^{-1}. For each of the angular approximations (P_1, and P_3), the magnitudes of both the scattering and absorption coefficients of the inhomogeneous ring indicate that a void is present but are overestimated. The P_3 result also provides a significantly improved reconstruction of the internal scattering region when compared to the P_1 angular approximation results. During both P_1 and P_3 angular approximations artefacts exist for both scattering and absorption reconstructions. These are manifested in the scattering coefficient peaks and have scattering coefficient extending up to 24.0cm^{-1}.

Figure 6.25 depicts absorption coefficient (left hand graphs) and scattering coefficient (right hand graphs) as a function of distance along the x-axis for the P_3 and P_1 angular

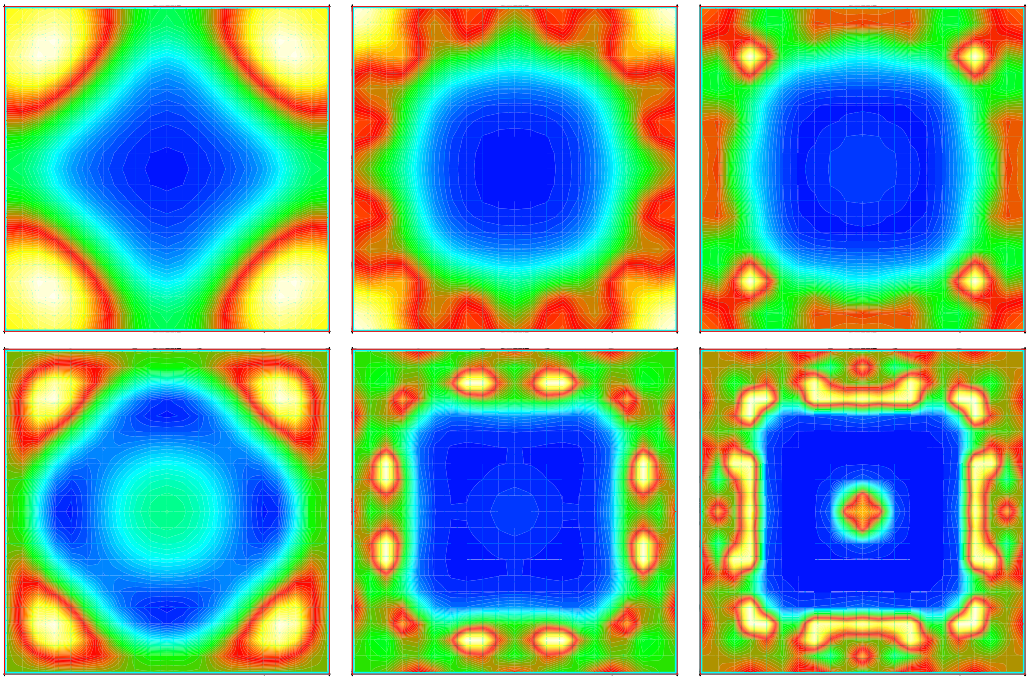


Figure 6.23: The absorption (top set of images) and scattering (bottom set of images) reconstruction results after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively, for the problem geometry described in figure 6.22 and table 6.5. A P_3 angular approximation was used to perform the transient inversion calculations.

approximations described in figures 6.23, and 6.24 respectively. A line is drawn from the middle of the y-axis, at co-ordinates (0.0, 2.5), (5.0, 2.5). The dotted line represents the exact solution. The top two graphs illustrate absorption and scattering coefficient as a function of distance after the first iteration. The middle graphs illustrate absorption and scattering coefficient as a function of distance after the tenth iteration and the bottom graphs illustrate absorption and scattering coefficient as a function of distance after the fiftieth iteration.

The absorption coefficient reconstructions illustrated in figure 6.25 show that the void is largely resolved somewhere between the first and tenth iteration. After 50 iterations the absorption coefficient extends from approximately $\mu_a = 0.07\text{cm}^{-1}$ to approximately $\mu_a = 0.01\text{cm}^{-1}$. Figure 6.25 also illustrates that even, after 50 iterations, the ring is never reconstructed for the P_1 angular approximation, while the ring is partially resolved for the P_3 angular approximation.

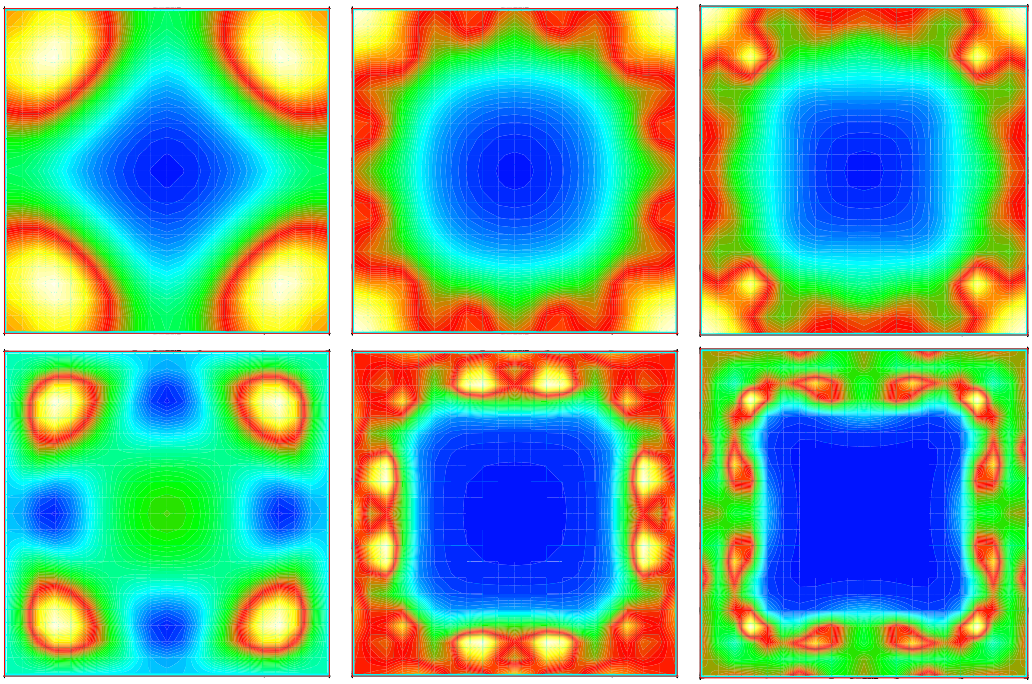


Figure 6.24: The absorption (top set of images) and scattering (bottom set of images) reconstruction results after (from left to right) 1 iteration, 10 iterations and 50 iterations respectively, for the problem geometry described in figure 6.22 and table 6.5. A P_1 angular approximation was used to perform the transient inversion calculations.

Like the absorption coefficient reconstruction, scattering coefficient reconstruction results resolve the void-like region somewhere between the first and tenth iteration. The tenth and fiftieth iteration graphs illustrate the ring-like structure emerging for the P_3 angular approximation results which is illustrated in figure 6.25 as a peak that extends between a distance of $x = 2$cm and $x = 3$cm. The presence of the ring structure is confirmed in the spatial reconstruction illustrated in figure 6.23. After 50 iterations the ring is fully reconstructed and has the correct approximate thickness and scattering coefficient magnitude. Two of the image artefacts are also reconstructed in figure 6.25. The peaks have scattering coefficients extending to approximately $\mu_a = 21\text{cm}^{-1}$.

Similar P_1 and P_3 calculations were performed for steady state radiation transport, the results of which are illustrated in [43]. While both steady state and time dependent radiation transport fail to reproduce the ring-like structure for P_1 angular expansions the visualisations obtained using time dependent information appear sharper and more accurate than

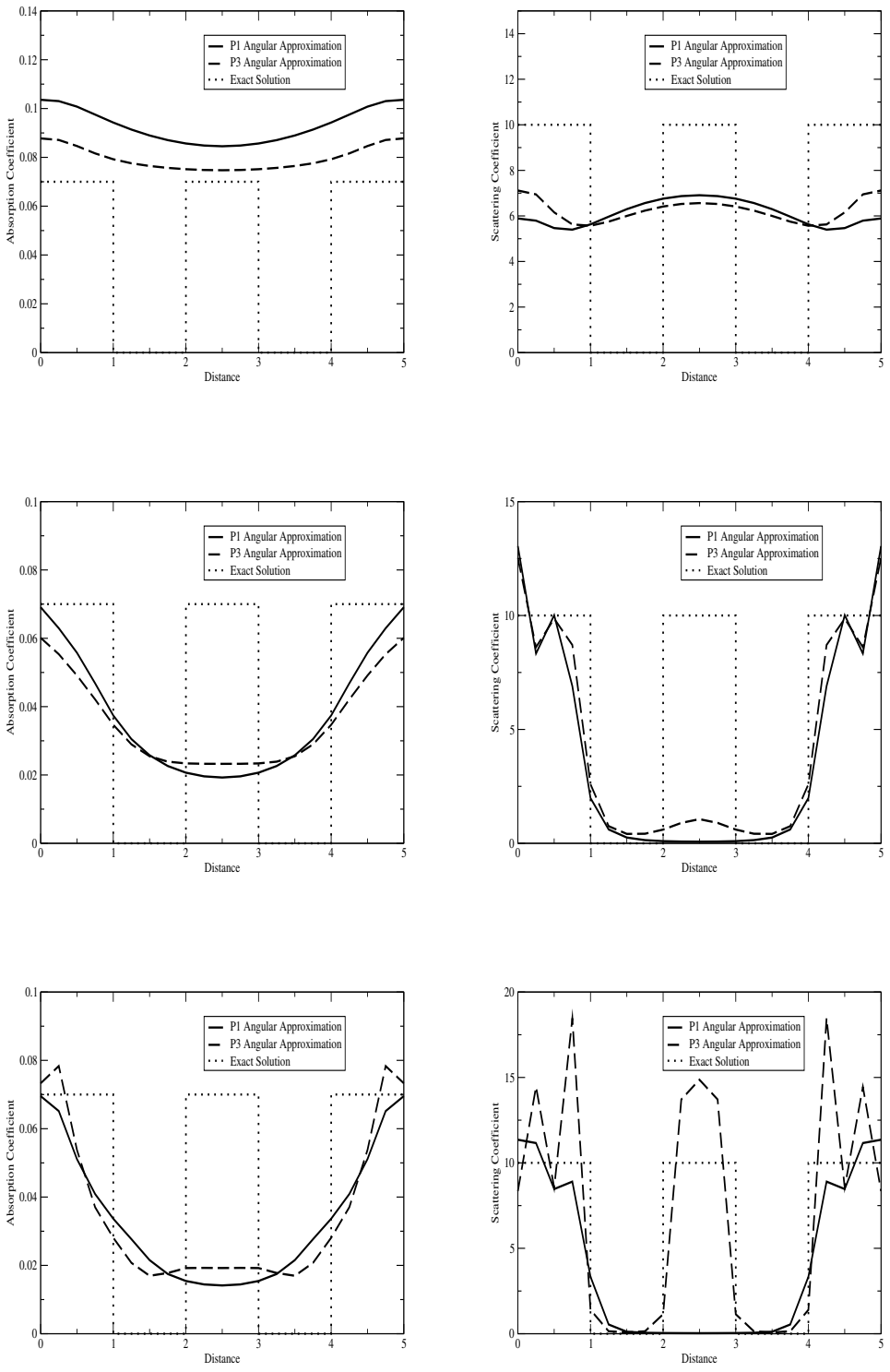


Figure 6.25: Absorption coefficient (left hand graphs) and scattering coefficient (right hand graphs) as a function of distance along the x-axis -for the transient P_3 and P_1 angular approximations described in figures 6.23, and 6.24- after 1 iteration (top graphs), 10 iterations (middle graphs) and 50 iterations (bottom graphs) respectively. The dotted lines represent the exact solution

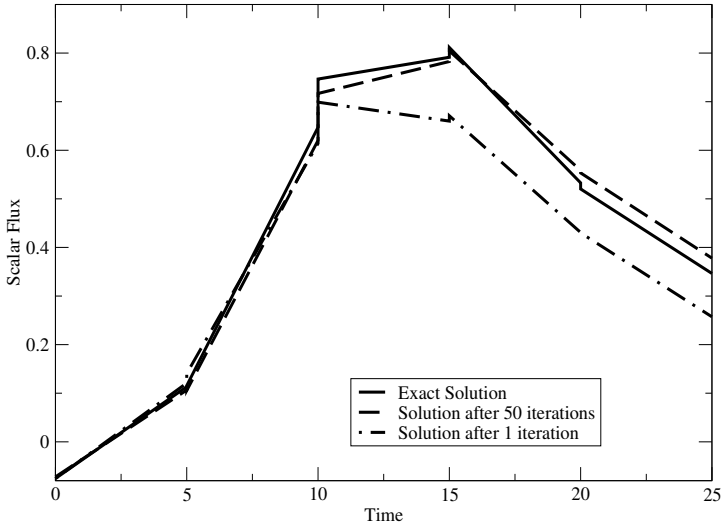


Figure 6.26: Scalar flux versus time for the first source and the third detector position. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (long dashed line), and the forward model result after 50 iterations (dot-dashed line). The solution approaches convergence as the number of iterations is increased.

those obtained using steady state radiation transport. During P_3 angular approximations, both time dependent and steady state radiation transport calculations resolve the transparent ring during scattering coefficient reconstructions. The ring appears clearer and more accurately resolved when transient information is used. Time dependent radiation transport also goes some way to resolving the transparent ring during P_3 absorption reconstructions. The transparent ring was never resolved for absorption reconstructions during steady state radiation transport calculations (even when a P_5 angular expansion is considered).

Figure 6.26 depicts scalar flux as a function of time at the first source and the third detector position. The x- and y- position co-ordinates of each individual source and detector are the same as the x- and y- position co-ordinates of the source and detectors described in experiment 2 and are illustrated in table 6.1. The first source is placed at coordinates (1.0, 0.0) and the third detector is placed at coordinates (3.0, 0.0). The graph compares the

exact solution (solid line) to the forward model result after 1 iteration (long dashed line), and the forward model result after 50 iterations (dot-dashed line). A marked improvement in the convergence of the scalar flux (at the third detector) is illustrated as the number of iterations is increased to 50. During the first 50 iterations of the P_1 and P_3 angular expansions, the error functional is reduced by more than two orders of magnitude. The P_1 angular expansion reduces the error functional from $F = 9987$ to $F = 160.0$ and the P_3 angular expansion reduces the error functional from $F = 7540$ to $F = 68.42$.

6.6 Conclusion

A multi-dimensional inversion scheme has been developed to describe the transport of near infrared photons during medical optical tomography. The main contribution of this chapter is the development of a time dependent inversion scheme. A series of applications have been considered to demonstrate the enhanced imaging achievable using the time dependent form of the Boltzmann transport equation. Our approach uses a Discontinuous Galerkin discretisation of the time domain, which has the ability to use large time steps while still achieving a high order of accuracy in time. The discretisation has been designed specifically for inverse problems and is designed to work well for both optically thick (diffusive) and transparent media.

Four synthetic problems have been used to demonstrate the simultaneous reconstruction of absorption and scattering material properties inside a host medium. While, the phantom problems are simple it is hoped that the inversion method will eventually be applied to real life medical imaging scenarios. The examples illustrate a significant improvement in scattering and absorption coefficient reconstructions when time dependent information is incorporated into the imaging domain.

Chapter 7

Optical Imaging with Voids using Transport Theory

Contents

A multi-dimensional inversion technique has been developed to describe the transport of near infrared photons during medical optical tomography. The transport of optical radiation is described according to the one-speed Boltzmann transport equation. The strengths of the forward model include the ability to model both optically thick and transparent media and the ability to include transient information. A least squares functional optimisation method is used to solve the inverse problem and results in the reconstruction of the spatial distribution of scattering and absorption coefficients inside the imaging domain.

Using a full transport theory forward model to predict light propagation enables accurate modelling of transparent, void-like regions, which are commonly found in biological media. Examples of void-like regions include the cerebrospinal fluid that surrounds the human brain and synovial fluid contained between joints. This chapter compares model reconstruction results of media containing transparent regions using full transport theory to reconstructions performed using a P_1 approximation to the Boltzmann transport equation. Both steady state and time dependent reconstruction results are compared.

7.1 Introduction

The non-invasive properties of optical radiation together with optical tomography's relative affordability compared to most conventional imaging technology make it a desirable method to develop and utilise as a mainstream imaging procedure. In the field of biomedical optics [12], pathological conditions provide a contrast in tissue optical properties when compared to healthy tissue [14]. In the field of optical imaging, such contrasts are exploited to provide (three dimensional) qualitative and quantitative optical images.

Strongly scattering problems, such as those posed in near infrared optical tomography, are complex and non-linear. There is not a generally applicable direct method for resolving highly scattering phenomena, and instead, non-linear inversion methods are used to evaluate what optical properties would have created the original signal [3, 4, 26, 35, 40, 41, 42, 68, 69, 70]. The techniques described in chapters 4 and 5 enable the reconstruction of optical property distribution for arbitrary model geometries and for arbitrary source and detector positions along the boundary of the domain. The inversion is posed as a functional optimisation [41, 69, 74, 75, 76] problem. Functional optimisation techniques utilise a forward model that provides detector readings based on estimates of the distribution of optical properties inside the medium. The functional is minimised using a Levenberg-Marquardt optimisation scheme [71, 72, 73]. Iterations are performed until the predicted data agrees with the detector readings. The final distribution of optical properties is then displayed as an image.

Optical property reconstruction is strongly dependent on the accuracy of the forward model. If the forward model does not accurately describe photon propagation inside a given medium then the reconstruction scheme becomes erroneous. When considering radiation transport, many research groups use forward models that rely on the diffusion approximation to the generally applicable Boltzmann transport equation [3, 4, 26, 173]. While diffusion theory [32, 33, 34, 112] remains a good approximation for light transport in highly scattering media and is considered advantageous over transport theory because of the much shorter computation times required to perform an image reconstruction [174], diffusion theory fails for media or regions contained inside diffusive media that are not considered diffusive [36, 37, 175]. In other words, the diffusion approximation fails for media in which sources and detectors are not sufficiently far apart, media with small geometries where boundary effects dominate and for tissue structures, which do not exhibit diffusive behaviour. Diffusion theory specifically fails in highly absorbing regions, regions in which μ_a is comparable to or greater than μ_s, and regions in which μ_a and μ_s are both very low (void-like regions).

Although biological tissue approximates a diffusive medium, low-scattering regions exist that can violate the conditions needed to apply diffusion theory. For example highly scattering brain tissue is enclosed in a layer of virtually transparent cerebrospinal fluid (CSF) which circulates around the brain. CSF has an absorption coefficient and a scattering coefficient of approximately zero and therefore violates the conditions required to perform diffusion calculations. Cerebrospinal fluid is also found in ventricles located inside brain matter and in regions where the brain folds in on itself (sulci). In the detection and treatment of rheumatoid arthritis a similarly clear synovial fluid exists between joints, which also violates the scattering and absorption characteristics required to perform diffusion calculations. It is likely that the presence of CSF fluid in the human brain and synovial fluid contained between joints could cause detrimental effects to the propagation of light during near infrared optical tomography.

Regions that do not comply with the constraints of the diffusion approximation must be accounted for if optical tomography should develop into a conventional imaging tool. Numerical methods that can tolerate tissues that do not comply with the constraints of the diffusion approximation include Monte Carlo [38, 39] and full radiation transport models

that rely on the Boltzmann transport equation [40, 41, 42, 43, 44, 45, 46, 47, 48, 50, 51]. Although Monte Carlo methods provide correct solutions the simulations are often too slow to be used for realistic geometries such as those required in brain and breast imaging. Radiation transport models that solve the Boltzmann transport equation provide a suitable alternative to diffusion theory but are often difficult to solve and require much larger computational time scales to perform an image reconstruction. As such, many research groups still rely on either diffusion theory or radiation-diffusion hybrid models to perform tomographic image reconstructions [52, 53, 54]. Our model uses a finite element, Petrov-Galerkin discretisation of the Boltzmann transport equation [58, 59, 60, 61, 62]. Angular discretisation is achieved using spherical harmonics. Both steady state and time-dependent calculations are performed.

This chapter is organised as follows. Section 7.2 describes the diffusion approximation as a means of modelling photon transport, while the numerical applications described in section 7.3 demonstrate the feasibility of the proposed method and specifically investigates the applicability of using optical tomography to image the neonatal brain. Our findings are summarised in section 7.4.

7.2 The One-Speed Diffusion Approximation

The diffusion approximation [32, 33, 34, 112] represents an approximation of the spherical harmonic, angular flux expansion of the Boltzmann transport equation truncated to first order (i.e. a P_1 angular expansion). For a P_1 angular approximation, only the first two terms of the angular expansion are retained ($l = 0$ and $l = 1$). The final result yields:

$$\frac{1}{v}\frac{\partial \psi(\mathbf{r},t)}{\partial t} + \nabla \mathscr{J}(\mathbf{r},t) + \mu_a(\mathbf{r})\psi(\mathbf{r},t) = s(\mathbf{r},t), \tag{7.1}$$

$$\frac{1}{v}\frac{\partial \mathscr{J}(\mathbf{r},t)}{\partial t} + \frac{1}{3}\nabla\psi(\mathbf{r},t) + [\mu_a(\mathbf{r}) + \mu_s^{'}(\mathbf{r})]\mathscr{J}(\mathbf{r},t) = s_1(\mathbf{r},t). \tag{7.2}$$

In which, $\mu_s^{'}(\mathbf{r})$ is the reduced scattering coefficient and $\mathscr{J}(\mathbf{r},t)$ represents the photon

current. Two further approximations are introduced to simplify equations 7.1 and 7.2 further. Firstly, it assumed that the source term, $s(\mathbf{r}, \mathbf{\Omega}, t)$, is isotropic. This implies that the source term $s_1(\mathbf{r}, t)$ vanishes in equation 7.2. Secondly, it is assumed that the time derivative in equation 7.2 is also 0. That is, the rate of time variation of the current density is much slower than the collisional frequency, $v\mu_t(\mathbf{r})$. Re-arranging equation 7.2 yields an expression for the photon current density, $\mathscr{J}(\mathbf{r}, t)$, described in terms of the angular flux $\psi(\mathbf{r}, t)$: $\mathscr{J}(\mathbf{r}, t) = -\frac{1}{3[\mu_a(\mathbf{r}) + \mu_s'(\mathbf{r})]}\nabla\psi(\mathbf{r}, t)$.

If the photon diffusion coefficient, $D(\mathbf{r})$, is defined as: $D(\mathbf{r}) = \frac{1}{3[\mu_a(\mathbf{r}) + \mu_s'(\mathbf{r})]}$, then the one-speed diffusion approximation results:

$$\frac{1}{v}\frac{\partial\psi(\mathbf{r}, t)}{\partial t} - \nabla \cdot D(\mathbf{r})\nabla\psi(\mathbf{r}, t) + \mu_a(\mathbf{r})\psi(\mathbf{r}, t) = s(\mathbf{r}, t). \tag{7.3}$$

The diffusion approximation is a consequence of four assumptions [33]. That is, 1) the angular flux must be adequately described by its linearly anisotropic angular dependence, 2) the sources must be isotropic, 3) a one-speed approximation must be considered and 4) the photon current density changes slowly on a time scale compared to the mean collision time. The assumption of weak angular dependence is most commonly violated in medical optics. Weak angular dependence is often violated near boundaries or where material properties change dramatically from point to point over distances comparable to the mean free path, near localised sources, and in strongly absorbing media.

A P_1 angular approximation has been considered to compare diffusion and transport theory. However, the P_1 angular expansion is not completely equivalent to the diffusion theory described in equation 7.3. The diffusion approximation represents a hyperbolic system whereas a P_1 angular expansion represents a parabolic system. This affects the causality of the model: the P_1 model should have a non-zero solution at some distance for a given time during which photons have travelled at the speed of light (for a given model). The diffusion approximation gives a non-zero value instantaneously at all points in space which is nonphysical.

	Experiment 1		Problems 1 and 2		Problem 3	
Source/Detector Number	x	y	x	y	x	y
1	1.0	0.0	-2.5	0.0	0.0	3.5
2	2.0	0.0	-1.75	1.75	2.5	2.5
3	3.0	0.0	0.0	2.5	3.5	0.0
4	4.0	0.0	1.75	1.75	2.5	-2.5
5	5.0	1.0	2.5	0.0	0.0	-3.5
6	5.0	2.0	1.75	-1.75	-2.5	-2.5
7	5.0	3.0	0.0	-2.5	-3.5	0.0
8	5.0	4.0	-1.75	-1.75	-2.5	2.5
9	4.0	5.0				
10	3.0	5.0				
11	2.0	5.0				
12	1.0	5.0				
13	0.0	4.0				
14	0.0	3.0				
15	0.0	2.0				
16	0.0	1.0				

Table 7.1: The x co-ordinate and y co-ordinate source locations and detector locations for each of the inversion experiments described in chapter 7.

7.3 Numerical Examples

Presented are the simultaneous reconstruction results of absorption and scattering coefficients contained inside a given isotropic medium from measurements taken at the surface of the domain. Circular and square test objects containing embedded void-like regions are utilised as a crude representation of the human head or breast. The forward models used in the reconstruction of each of the test cases were simulated using the finite element representation of the Boltzmann transport equation described in the theory. Reconstructions

performed using full radiation transport theory (i.e. P_3 and P_5 angular approximations) are compared along with reconstructions performed using the diffusion approximation to the Boltzmann transport equation (i.e. P_1 angular approximations) to illustrate the need for using full transport theory when modelling voids. Steady state radiation transport is also compared alongside time dependent radiation transport to demonstrate the enhanced imaging achievable using the time dependent form of the Boltzmann transport equation. Highly resolved meshes are used during both forward and adjoint calculations to ensure good quality data and the same sized mesh is used to generate both the synthetic data and the inversion reconstruction. As an initial guess, for each of the problems considered, a homogeneous medium is chosen with optical properties equal to those of the background media. The various locations of the individual sources and detectors for each of the subsequent experiments are described in table 7.1.

7.3.1 Experiment 1

The first simulation geometry is a square domain containing three embedded void regions. The domain is illustrated in figure 7.1 and is of size 5cm $\times$ 5cm. The three embedded inhomogeneities are each of size 1cm $\times$ 1cm. The absorption and scattering material properties of both the background medium and the embedded inhomogeneities are described in table 7.2. Sixteen equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward model and reconstruction model results. P_1 and P_3 angular approximations are used to obtain steady state forward and inversion calculations along with a Levenberg-Marquardt optimisation technique. The finite element mesh has 882 nodes and 1280 bi-linear quadrilateral elements. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. The x- and y- co-ordinates of each of the source and detector locations is illustrated in table 7.1.

Figure 7.2 illustrates scalar flux as a function of x- and y- spatial co-ordinates for the forward solution (top left hand side graph), the log of the scalar flux forward solution (top right hand side graph) and the adjoint solution (bottom graph), for a single source-detector

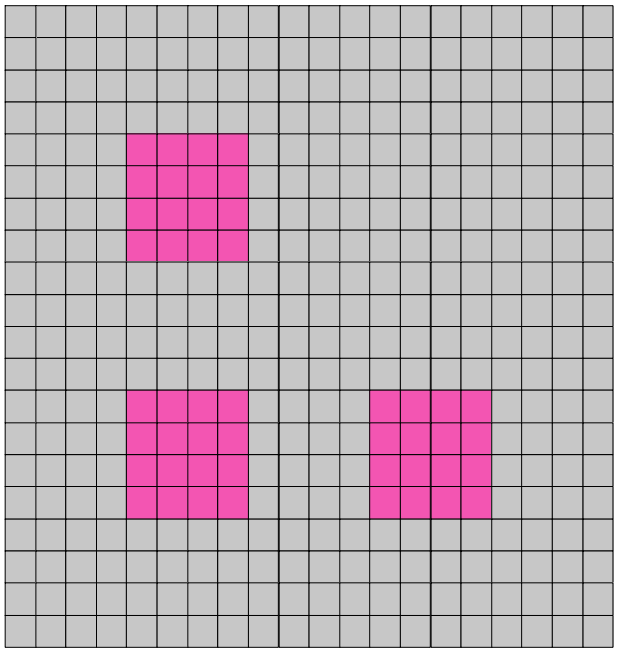


Figure 7.1: A two dimensional finite element square mesh for a phantom of size 5cm $\times$ 5cm. Three embedded voids are located 1cm deep inside the medium. The steady state forward model and inversion model reconstruction has 882 nodes and 1280 bi-linear quadrilateral elements.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.07	10.0
Top Left Inhomogeneity	0.0	0.0
Bottom Left Inhomogeneity	0.0	0.0
Bottom Right Inhomogeneity	0.0	0.0

Table 7.2: Optical absorption and scattering coefficients for a square domain of size 5cm $\times$ 5cm. Three embedded voids are contained inside the domain. The scattering and absorption material properties of the background medium and the respective voids are described.

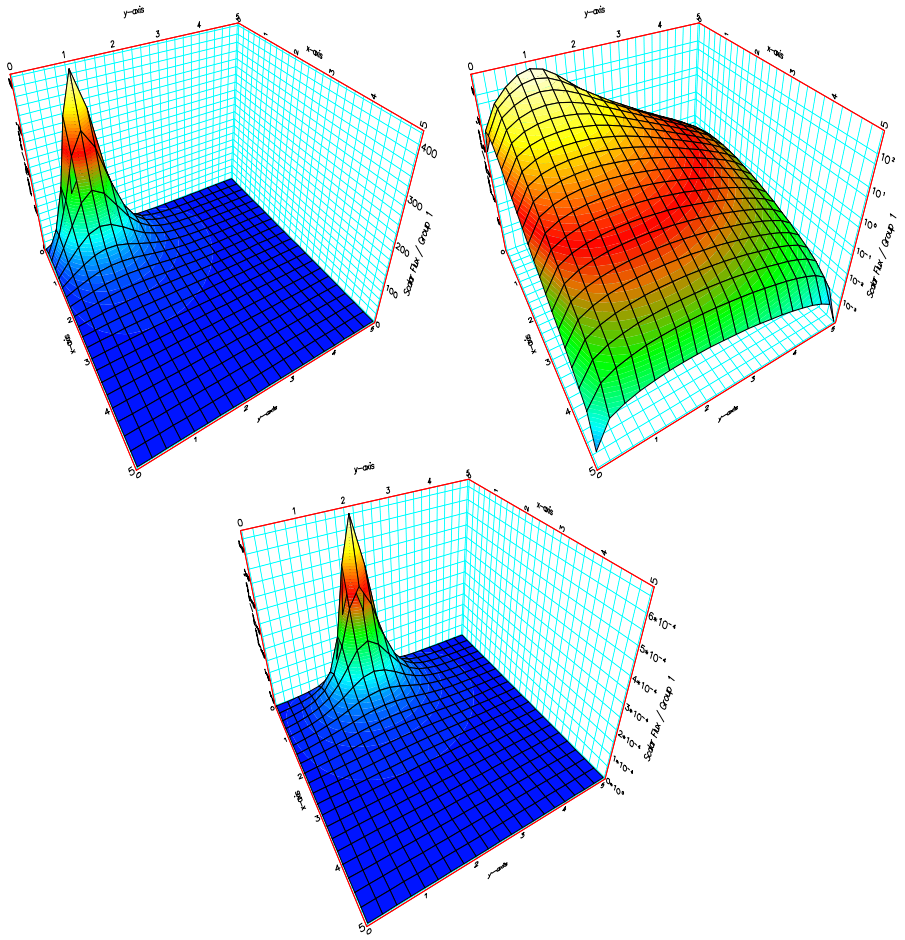
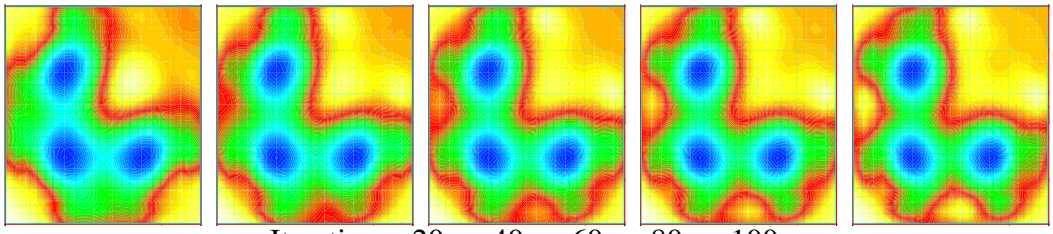


Figure 7.2: Scalar flux as a function of two dimensional space (x and y) for the scalar flux forward solution (top left hand side graph), the log of the scalar flux forward solution (top right hand side graph), and the scalar flux adjoint solution (bottom graph). A P_3 angular approximation is used to obtain each of the above images.

arrangement, using a P_3 angular expansion. The peaks illustrated in the forward solution diagrams indicate the position of the first isotropic source (at co-ordinates 0.0, 1.0) and how the scalar flux propagates through the medium. The adjoint solution represents the data misfit between the exact solution and the model prediction. The adjoint solution illustrates the adjoint flux obtained from the second detector, placed at co-ordinates (0.0, 2.0). The first and second order gradient of the error functional attempt to adjust the absorption and scattering material properties so that the forward model can better approximate the exact solution. As a result the scalar flux of the adjoint solution, the error functional, and the gradient of the error functional all tend to zero as better approximations to the exact solution are made.

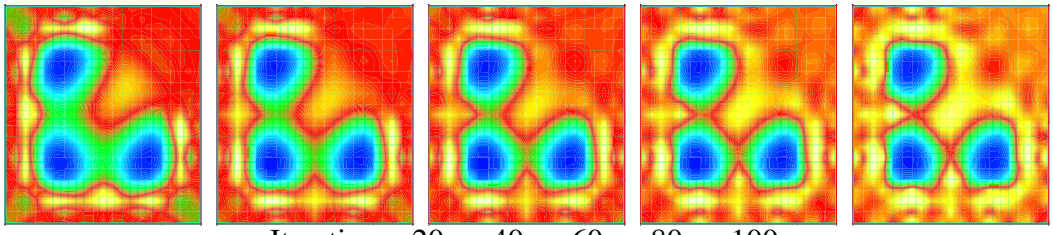
Figure 7.3 compares the absorption (top set of images) and scattering (bottom set of images) reconstruction results for the problem geometry described in figure 7.1 and table 7.2. A P_3 angular approximation was used to perform the inversion calculations. Figure 7.3 provides visual evidence that during the P_3 inversion calculation, the algorithm finds the correct location and approximate size of the embedded voids for both the absorption and scattering reconstructions. The relative magnitudes of the absorption and scattering coefficients are also approximated fairly accurately, although the three voids are never completely resolved. Table 7.3 illustrates the approximate magnitude ranges of the absorption and scattering coefficients after 20, 40, 60, 80 and 100 iterations respectively. The magnitude of the absorption coefficient is reduced by almost 50% during early stages of the inversion but only small changes are observed after twenty iterations. In contrast, after 100 iterations, the scattering coefficient associated with the embedded voids is reduced by almost two orders of magnitude. While the absolute magnitude of the scattering coefficients are most accurate after 100 iterations the scattering image is quite noisy compared to earlier iteration images (see figure 7.3). Performing calculations with larger initial regularisation penalties could improve image quality.

The error functional is reduced by more than three orders during the first 100 Levenberg-Marquardt iterations. The functional is reduced from $F = 3774.00$ to $F = 1.11$. The majority of the reconstruction is achieved during the first twenty iterations. This is because, firstly all three void-like regions are resolved to fairly accurate absorption and scattering coefficient magnitudes and secondly, the error functional undergoes its biggest reduction



Iteration = 20 → 40 → 60 → 80 → 100

Absorption Coefficient Images



Iteration = 20 → 40 → 60 → 80 → 100

Scattering Coefficient Images

Figure 7.3: Absorption (top set of images) and scattering (bottom set of images) reconstruction results during the first 100 iterations of a square domain problem containing three embedded voids. The inversion was performed using a P_3 angular approximation.

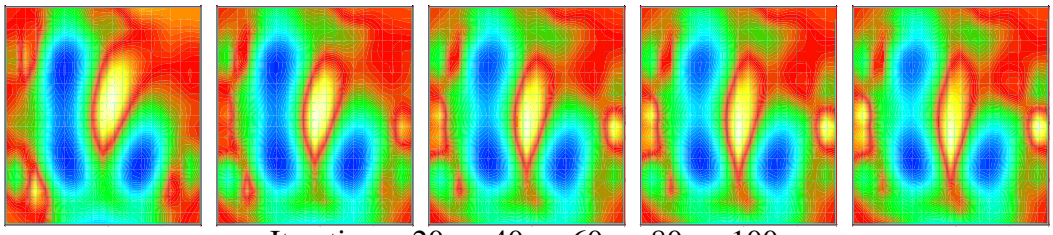
Iteration Number	$\mu_a[cm^{-1}]$ range	$\mu_s^{'}[cm^{-1}]$ range
20	0.0774 - 0.0436	15.7 - 1.39
40	0.0773 - 0.0426	15.3 - 1.08
60	0.0773 - 0.0423	14.9 - 0.964
80	0.0773 - 0.0421	14.2 - 0.859
100	0.0773 - 0.0416	13.7 - 0.750

Table 7.3: The range in optical absorption and scattering coefficients for the square domain described in figure 7.1. The lower limits correspond to the resolved absorption and scattering coefficients of the three void-like regions. Calculations were performed using a P_3 angular expansion.

(from $F = 3774.00$ to $F = 8.09$) during the first twenty Levenberg-Marquardt iterations. The absolute magnitudes of the absorption and scattering coefficients are still reduced during subsequent iterations although little is gained from using more than twenty.

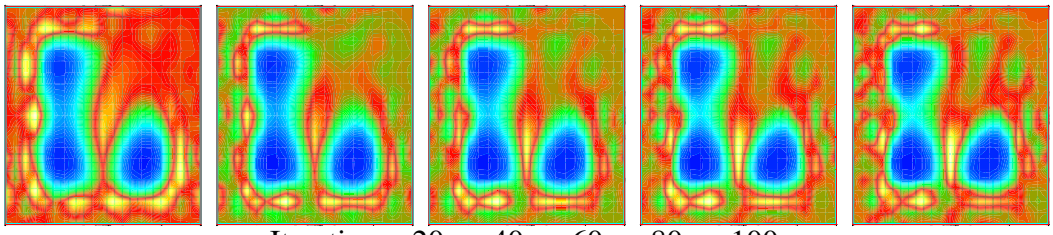
When considering radiation transport, many research groups use forward models that rely on the diffusion approximation to the generally applicable Boltzmann transport equation. The diffusion approximation amounts to a P_1 angular expansion. While diffusion theory remains a good approximation for light transport in highly scattering media, such as biological tissue, and is considered advantageous over transport theory because of the much shorter computation times required to perform image reconstructions, diffusion theory fails in regions that are not regarded as diffuse. In other words, the diffusion approximation fails when sources and detectors are not sufficiently far apart, for media with small geometries where boundary effects dominate and for tissue structures which do not exhibit diffusive behaviour. Diffusion theory specifically fails in highly absorbing regions, regions in which μ_a is comparable or greater than μ_s, and regions in which μ_a and μ_s are both very low. The transparent regions described in the theory above violate the conditions in which diffusion theory should be applied.

Figure 7.4 compares the absorption (top set of images) and scattering (bottom set of images) coefficient reconstruction results for the problem geometry described in figure 7.1 and table 7.2. A P_1 (diffusion theory) angular approximation was used to perform the inversion calculation. Figure 7.4 finds the correct location and approximate size of the three embedded voids for both absorption and scattering reconstructions, however cross-talk exists between the three voids. A higher degree of noise is also observed when compared to calculations performed using a P_3 angular approximation. The relative magnitudes of the absorption and scattering coefficients are approximated accurately, although the three voids are never completely resolved. Table 7.4 illustrates the approximate magnitude ranges of the absorption and scattering coefficients, during the P_1 angular expansion, after 20, 40, 60, 80 and 100 iterations respectively. The magnitude of the absorption coefficient is reduced by 50% during early stages of the inversion, while the scattering coefficient associated with the embedded voids is reduced by almost two orders of magnitude. Similar magnitudes were observed during the P_3 angular expansion.



Iteration = 20 → 40 → 60 → 80 → 100

Absorption Coefficient Images



Iteration = 20 → 40 → 60 → 80 → 100

Scattering Coefficient Images

Figure 7.4: Absorption (top set of images) and scattering (bottom set of images) coefficient reconstruction results during the first 100 iterations of a square domain problem containing three embedded voids. The inversion was performed using a P_1 angular approximation.

Iteration Number	$\mu_a[cm^{-1}]$ range	$\mu_s^{'}[cm^{-1}]$ range
20	0.0894 - 0.0402	16.5 - 1.06
40	0.0950 - 0.0379	19.6 - 0.726
60	0.0970 - 0.0382	19.4 - 0.539
80	0.0985 - 0.0355	18.8 - 0.476
100	0.0973 - 0.0348	18.8 - 0.419

Table 7.4: The range of optical absorption and scattering coefficients for the square domain described in figure 7.1. The lower limits correspond to the resolved absorption and scattering coefficients of the three void-like regions. Calculations were performed using a P_1 angular expansion.

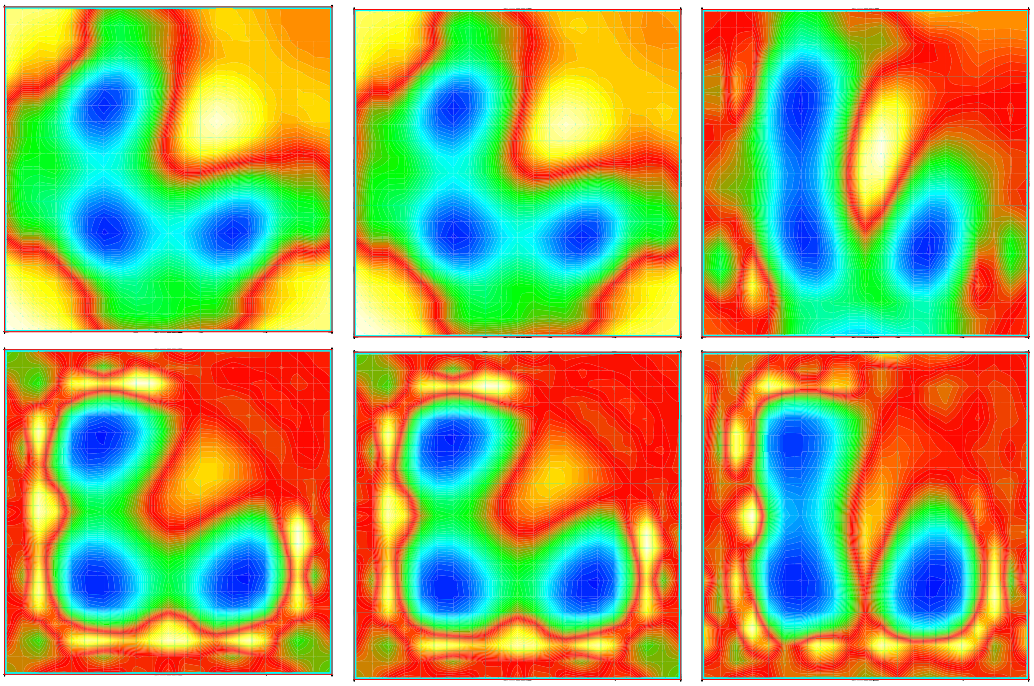


Figure 7.5: The absorption coefficient (top set of images) and scattering coefficient (bottom set of images) reconstruction results during P_5 (left hand side images), P_3 (middle images), and P_1 (right hand side images) angular expansions. The respective reconstructions were run for 20 iterations.

The error functional associated with the P_1 angular expansion is reduced by a similar magnitude compared to the first 100 Levenberg-Marquardt iterations of the P_3 angular expansion. The functional is reduced from $F = 5920.00$ to $F = 3.35$. Although the majority of the reconstruction is performed during the first 20 iterations, the P_1 angular expansion takes much longer to distinguish three distinct inhomogeneities. In fact the top left and bottom left inhomogeneities are never completely separated during the P_1 inversion for both scattering and absorption reconstructions. This is problematic when considering that the most important aspect of medical optical tomography is the ability to visualise the internal structure of a given medium. After 20 iterations the error functional is reduced from a magnitude of $F = 5920.00$ to a magnitude of $F = 45.25$. It is clear that a P_3 angular approximation provides a more accurate means of modelling radiation transport in void-like regions when compared to a P_1 angular approximation.

Figure 7.5 compares the absorption coefficient (top set of images) and scattering coeffi-

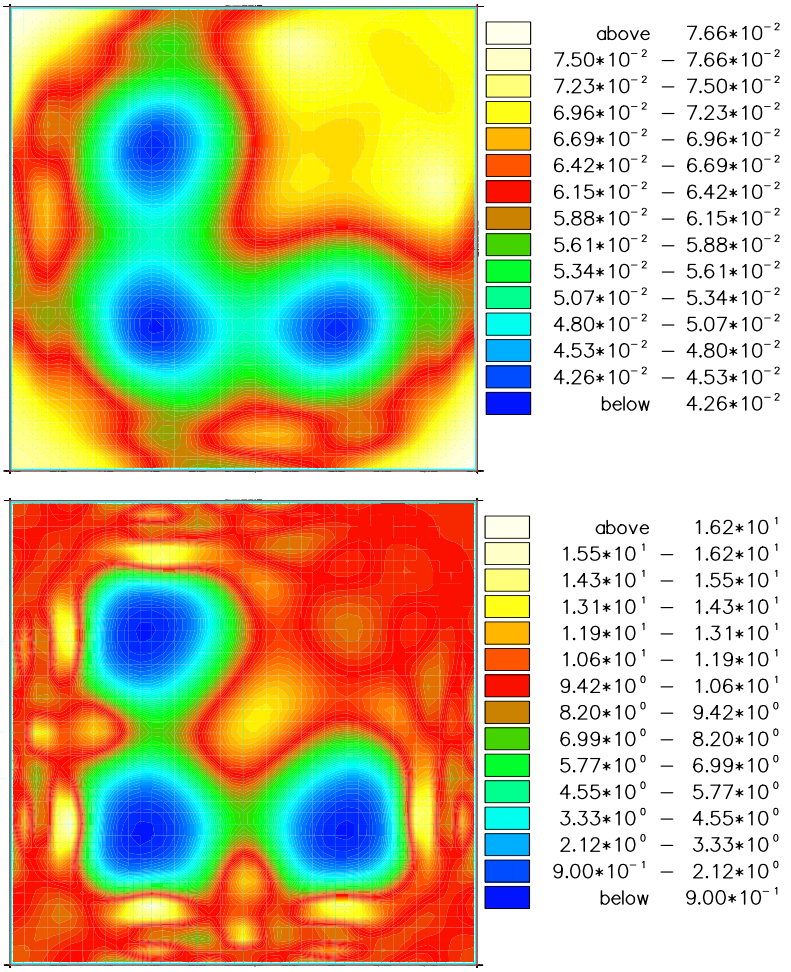


Figure 7.6: Absorption (top graph) and scattering (bottom graph) coefficient reconstruction results for a square domain problem using a P_3 angular expansion. The inversion was performed for 100 iterations. The problem geometry is the same as described in figure 7.3 but a refined finite element mesh is used.

cient (bottom set of images) reconstruction results during the first 20 iterations of a P_5 (left hand side images), P_3 (middle images), and P_1 (right hand side images) angular expansion. Both the P_5 and P_3 angular approximations resolve three specific inhomogeneities during both absorption and scattering reconstructions after 20 iterations. The P_1 angular approximation fails to resolve three specific homogeneities. A slight improvement in image quality is observed when comparing P_5 to P_3 angular expansions, however this is not significant given the increased computational effort required to perform P_5 calculations.

Figure 7.6 compares absorption (top graph) and scattering (bottom graph) coefficient re-

construction results for the problem geometry described in figure 7.1 and table 7.2. Calculations were performed using a P_3 angular expansion and a Levenberg-Marquardt optimisation technique. While the problem geometry and optimisation conditions remain largely the same as the conditions described in figure 7.3, the mesh used for the forward calculation and the inversion reconstruction was much more refined. The new uniform finite element mesh has 3362 nodes and 4960 bi-linear quadrilateral elements. The absorption and scattering penalty levels (γ_a and γ_{s0}) are also increased to initial values of 10,000 before being annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. It is necessary to increase the regularisation penalty level because refined meshes with more material property unknowns are more sensitive to the effects of ill-posedness. Figure 7.6 illustrates that little difference is observed in both the absorption and scattering coefficient reconstruction results when a finite element mesh with more nodes is considered. The only noticeable difference appears to be a slightly sharper image. The reconstruction reduced the error functional from a magnitude of $F = 3966$ to a magnitude of $F = 1.25$ after 100 Levenberg-Marquardt iterations.

Incorporating Noise into the Steady State Imaging Domain

It is necessary to study the effect of measurement error on image reconstruction. A pseudo random number generator has been used to add independent Gaussian noise to each measurement. The independent identical Gaussian noise provides a good approximation to the true statistics of the data error [176]. In order to generate a data set containing errors, the angular flux $\Psi_{s\,\tau\,\mu\,i}$, detected from the forward calculation is multiplied by a random error term $E_{s\,\tau\,\mu\,i}(\alpha)$,

$$\Psi_{s\,\tau\,\mu\,i} \rightarrow (1 + E_{s\,\tau\,\mu\,i}(\alpha))\Psi_{s\,\tau\,\mu\,i}. \tag{7.4}$$

The random error is generated from a Gaussian distribution with mean zero and a standard deviation σ_{sd}. For each angular moment, a random number α is generated where α has equal probability of being any real number in the inteval $[0, 1]$. The error term is then calculated by finding the value $E_{s\,\tau\,\mu\,i}(\alpha)$, such that,

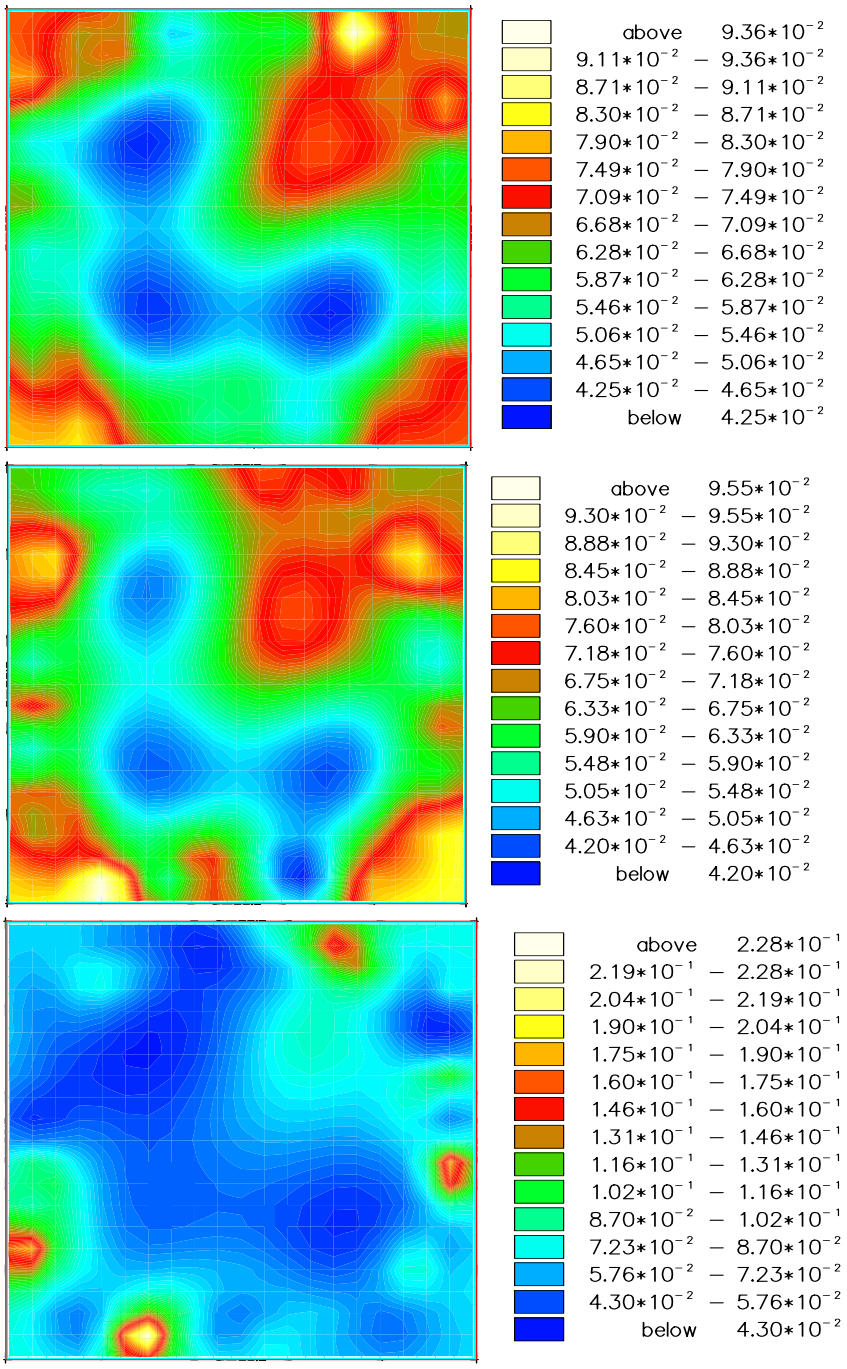


Figure 7.7: Absorption coefficient reconstruction results for a square domain problem containing three transparent voids. The inversion was performed for 20 iterations using a P_3 angular expansion. The images compare 5 $\%$ random noise (top image) contained in the data (top image), to reconstructions performed using 10 $\%$ random noise (middle image) and 20 $\%$ random noise (bottom image).

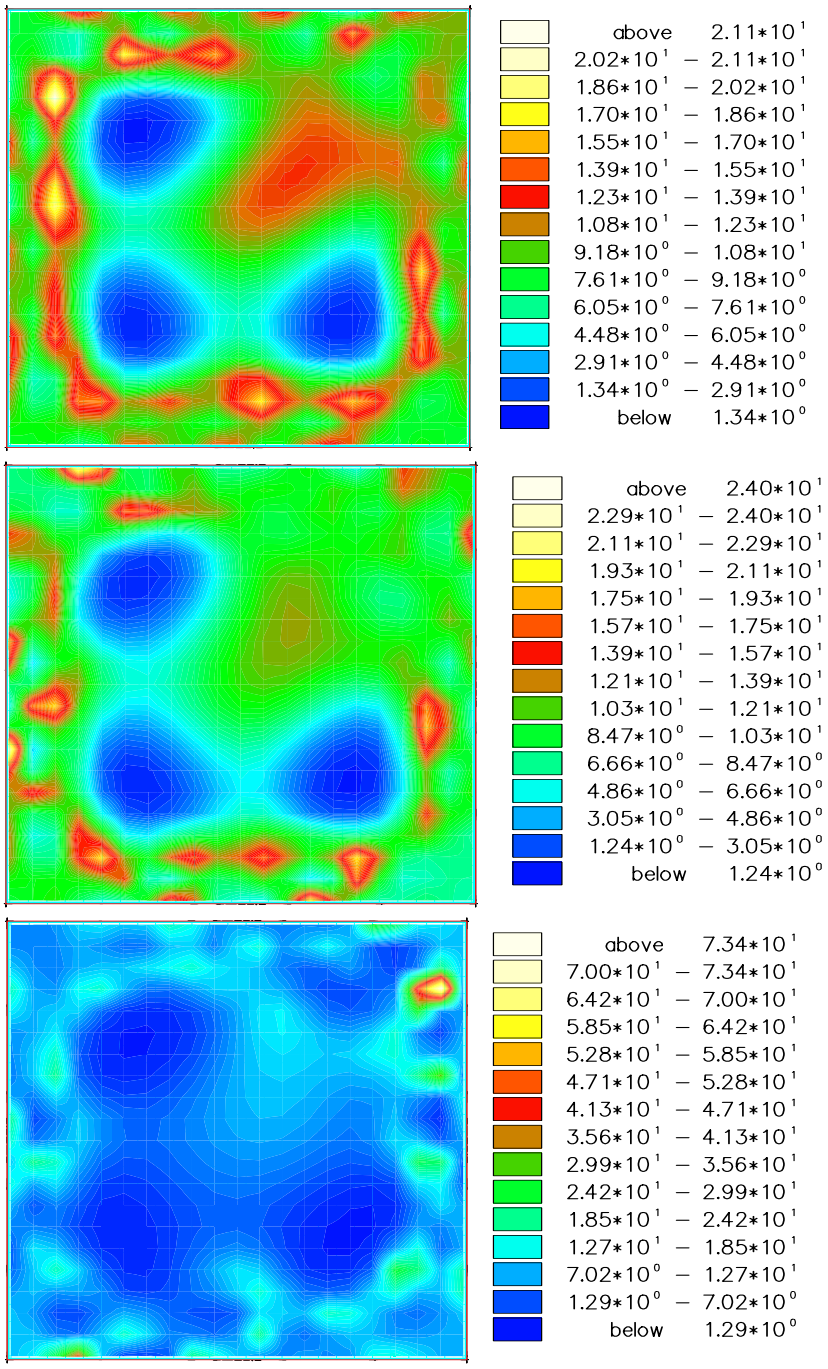


Figure 7.8: Scattering coefficient reconstruction results for a square domain problem containing three transparent voids. The inversion was performed for 20 iterations using a P_3 angular expansion. The images compare 5 $\%$ random noise (top image) contained in the data (top image), to reconstructions performed using 10 $\%$ random noise (middle image) and 20 $\%$ random noise (bottom image).

Tissue Type	Sample	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$	Reference
Brain of (24 week gest) neonate	in vitro	0.215	7.48	[133]
Brain of (28 week gest) neonate	in vitro	0.373	6.73	[133]
Neonate white matter (40 week gest)	in vitro	0.373	6.59	[133]
Neonate grey matter (40 week gest)	in vitro	0.460	5.29	[133]
Adult white matter	in vitro	0.06	85.0	[133]
Adult grey matter	in vitro	0.35	22.0	[133]
Cerebrospinal Fluid (CSF)		0.022	≈ 0	[4]
Pig skull	in vitro	0.25	18.0	[136]
Skin: Dermis	in vitro	0.13	20.0	[38]
Skin: Subdermis	in vitro	0.08	12.0	[38]

Table 7.5: Optical properties of various neonatal brain tissue at wavelength λ=800nm

$$\frac{1}{\sigma_{st}(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{E_s \tau \mu i(\alpha)} e^{\frac{-x^2}{\sigma_{st}^2}} dx = \alpha \tag{7.5}$$

The standard deviation determines the severity of the error in the data. In the following numerical examples the error data is defined as a percentage of the original singnal detected. The standard deviation is set as this set percentage, for example $5\% \rightarrow \sigma_{st} = 0.05$, so that when equation 7.4 is applied, the standard deviation in the error data has this specified value.

Figure 7.7 illustrates absorption coefficient reconstruction results while figure 7.8 illustrates scattering coefficient reconstruction results for the square domain problem containing three transparent voids. During both scattering and absorption reconstructions, the inversion was performed for 20 iterations using P_3 angular expansions. The images compare 5 $\%$ random noise (top images) contained in the data, to reconstructions performed using 10 $\%$ random noise (middle images) and 20 $\%$ random noise (bottom images). The absorption and scattering coefficients of the background medium and the embedded void are the same as the absorption and scattering coefficients described in table 7.2.

Pathological Condition	Characterised by	Caused by
Haematoma and Haemorrhage	A collection of blood inside the brain.	Ruptured blood vessels
Periventricular Haemorrhage (PVH)	A collection of blood inside the brain 's periventricular white matter	80% of periventricular haemorrhages are caused by ruptured capillaries in the brain's germinal layer.
Hypoxic Ischemia (HIE) and Birth Asphyxia	A region of low blood flow or a region of low blood oxygenation	Usually occurs because of an abnormal exchange between the placenta and foetus or if the umbilical cord is compressed during labour.

Table 7.6: Common pathological conditions suffered by both term and preterm neonates that could benefit from near infrared optical imaging.

The algorithm finds the correct location and approximate size of the embedded voids for both absorption and scattering reconstruction results when errors of 5 % and 10 % random noise are considered. When a substantial amount of noise (20 % random noise) is considered, the algorithm struggles to reconstruct the void regions for absorption coefficient reconstruction results. The presence of noise effects image resolution and contrast. As the magnitude of the random noise increases, image quality decreases. However, random statistical error rarely extends above a few percent, in which case, the 20 % random noise calculation becomes unphysical. The algorithm therefore copes fairly well with the addition of random noise. The magnitudes of both absorption and scattering coefficent reconstruction results are approximated fairly accurately.

7.3.2 Experiment 2 - A Crude Approximation of the Neonatal Brain

The neonatal brain provides a suitable model for optical imaging applications. MRI and CT scanners are costly and largely unsuitable in neonatal intensive care because of the need to transport a seriously ill child to the scanner. Optical tomography provides a bedside imaging system that can be used to provide both anatomical and physiological information about brain tissue. The neonatal brain also provides a simplified model of the adult brain. Characteristics of the neonatal brain that benefit deep tissue imaging are fourfold. Firstly neonates have smaller head sizes than adults. Secondly, the neonatal brain is surrounded by a soft, thin and flexible skull which deforms easily under light pressure. The skull bones are also not fully mineralised and the structures not fused, resulting in a lower average scattering coefficient compared to the adult skull. The CSF fluid that surrounds the human brain exhibits low absorption and almost no scattering and can therefore hamper deep tissue imaging by acting as a light guide. Neonates have a thinner layer of CSF fluid surrounding the brain, and are therefore less susceptible to such effects. Finally, the neonatal brain's grey and white matter regions are thought to have lower scattering coefficients and a comparatively small scattering coefficient mismatch. This benefits the penetration of near infrared light deep in to the brain's white matter, thus enabling deep imaging diagnostics.

It is necessary to consider a few limitations when considering neonatal optical tomography. Firstly the optical properties of the head, especially the neonatal head, are not very well known (some experimental data obtained from refereed journals is described in table 7.5). Further experimental research is crucial in order to create realistic brain phantoms. Secondly the neonatal brain and head change rapidly and grow enormously in size and complexity during pregnancy and soon after birth. This makes data acquisition problematic as a cap is placed onto the neonatal head to obtain measurements. If optical tomography is used in conjunction with MRI or CT, it may also be necessary to perform further images using these techniques. This is both costly and prohibitive and does not overcome the problem of having to transport the sick child to the imaging equipment. Common pathological conditions -suffered by neonates- that could benefit from near infrared optical imaging are illustrated in table 7.6.

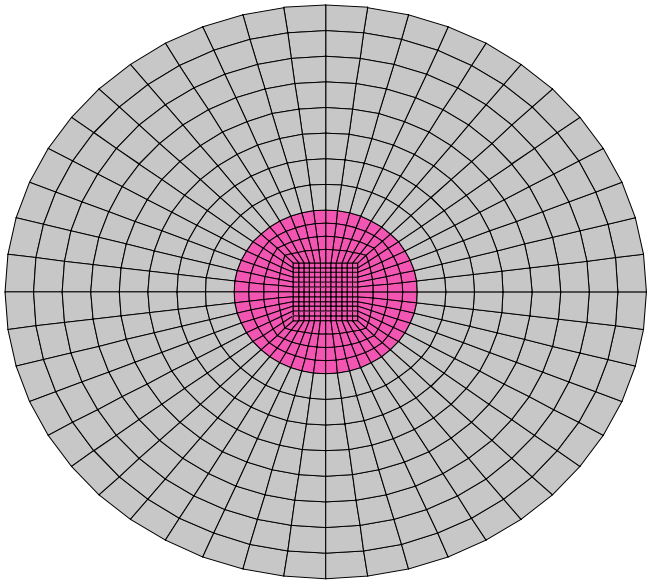


Figure 7.9: A two dimensional structured finite element circular mesh of diameter 5cm with an embedded void region of diameter 2cm located at the centre of the domain. The steady state problem has 1186 nodes and 1760 bi-linear quadrilateral elements.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.2	7.5
Inhomogeneity	0.0	0.0

Table 7.7: Optical absorption and scattering coefficients for a circle 5cm in diameter and a void placed at the centre of that circle 2cm in diameter. The background absorption and scattering coefficients approximate that of a neonatal brain.

Problem 1

The simulation geometry described in figure 7.9 depicts a circle 5cm in diameter containing a void situated in the centre of the circle that is 2cm in diameter. The absorption and scattering coefficients of the medium and the inhomogeneous void are contained in table 7.7. Eight equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward model and reconstruction model results. A P_3 angular approximation remains the most accurate means of modelling void

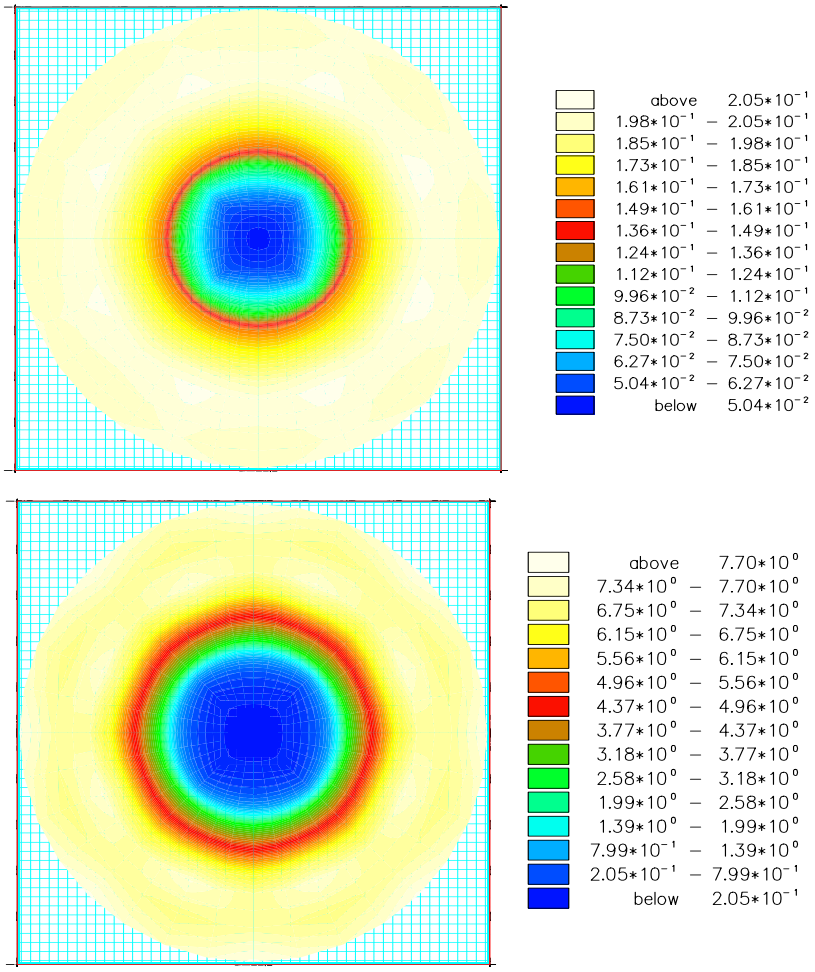


Figure 7.10: Absorption (top image) and scattering (bottom image) reconstruction results for the circle in circle geometry described in figure 7.9 and table 7.7. The inversion converged after 24 iterations using a P_3 angular expansion.

regions and is used to obtain the steady state forward and inversion calculations alongside a Levenberg-Marquardt optimisation method. The finite element mesh used for the forward and inverse reconstruction contained 1186 nodes and 1760 bi-linear quadrilateral elements. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. The x co-ordinate and y co-ordinates of each of the source and detector locations is illustrated in table 7.1. The optical absorption and scattering coefficients of the background medium approximate those of the brain of a 24 week gestation neonate [133].

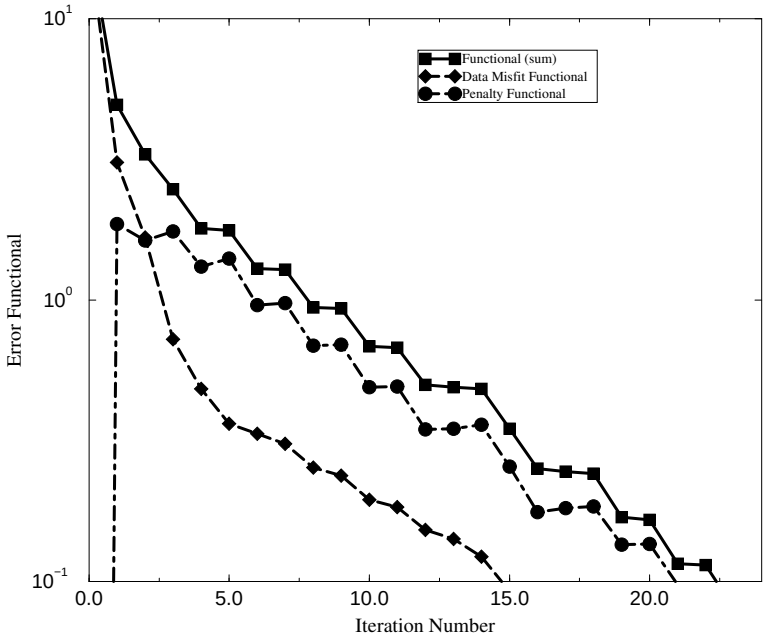


Figure 7.11: Error functional as a function of iteration number for the first 24 iterations of the P_3 angular approximation described in figure 7.9 and table 7.7. The error functional is reduced most dramatically during early stages of the inversion.

Figure 7.10 compares the converged absorption (top image) and scattering (bottom image) reconstruction results for the problem geometry described in figure 7.9 and table 7.7. Figure 7.10 provides visual evidence that during P_3 inversion calculations, the algorithm finds the correct location and approximate size of the cylindrical void for both the absorption and scattering reconstructions. The relative magnitudes of the absorption and scattering coefficients are also approximated fairly accurately. The resulting absorption coefficient, μ_a, has a range of 0.205cm^{-1} to 0.0505cm^{-1}, while the magnitude of the scattering coefficient is between 7.7cm^{-1} to 0.205cm^{-1}. The magnitudes of both the scattering and absorption coefficients of the inhomogeneity indicate that a void is present but are overestimated.

Figure 7.11 illustrates the error functional as a function of the iteration number during the first 24 iterations of the P_3 angular approximation described in figure 7.9 and table 7.7. The P_3 reconstruction converged after 24 iterations to an error functional of $F =$

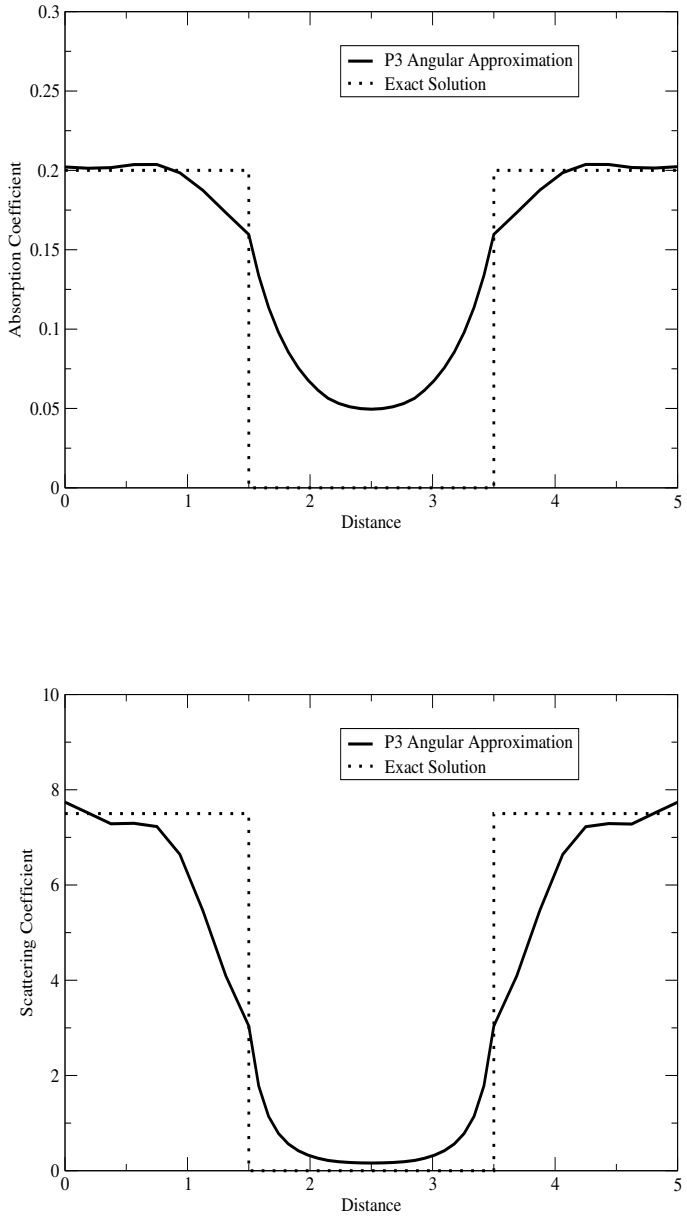


Figure 7.12: Absorption coefficient (top graph) as a function of distance along the x-axis and scattering coefficient (bottom graph) as a function of distance along the x-axis for the P_3 angular approximation described in figure 7.10. The inversion was performed for 24 iterations. A line is drawn trough the y-axis, at co-ordinates (-2.5, 0.0), (2.5,0.0). The dotted line represents the exact solution.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.2	7.5
Inhomogeneity 1	0.0	0.0
Inhomogeneity 2	0.0	0.0

Table 7.8: Optical absorption and scattering coefficients for a circle 5cm in diameter containing two irregular shaped void regions deeply embedded in the domain.

0.0797. The convergence criterion is chosen so that once convergence is met, increasing the number of iterations does not change the inversion model significantly. The error functional is reduced most dramatically during early stages of the inversion (during the first five iterations). Deeply embedded material structure emerges during iterations in which the penalty contribution to the error functional dominate.

Figure 7.12 depicts absorption coefficient (top graph) as a function of distance along the x-axis and scattering coefficient (bottom graph) as a function of distance along the x-axis, for the P_3 angular approximation described in figure 7.10. The dotted line in each of the graphs represents the exact solution. The inversion was performed for 24 iterations. A line is drawn in the middle of the y-axis, at co-ordinates (-2.5 0.0), (2.5, 0.0). The void region is clear in both the scattering and absorption reconstruction results and both the absorption coefficients and scattering coefficients are reduced markedly within the vicinity of the void. The size of the void is reconstructed most accurately for absorption reconstruction results and extends along a diameter of approximately 2cm along the x-axis. The scattering coefficient reconstruction produces an image that is slightly enlarged.

Problem 2

The simulation geometry described in figure 7.13 depicts a circle 5cm in diameter containing two irregular shaped, deeply embedded void regions. The absorption and scattering coefficients of the medium and the inhomogeneous voids are contained in table 7.8. The background scattering and absorption coefficient approximates that of the neonatal brain.

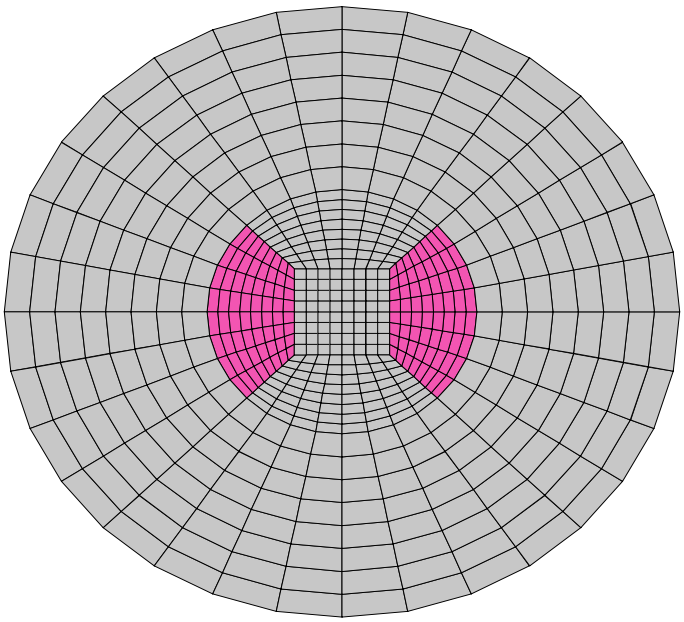
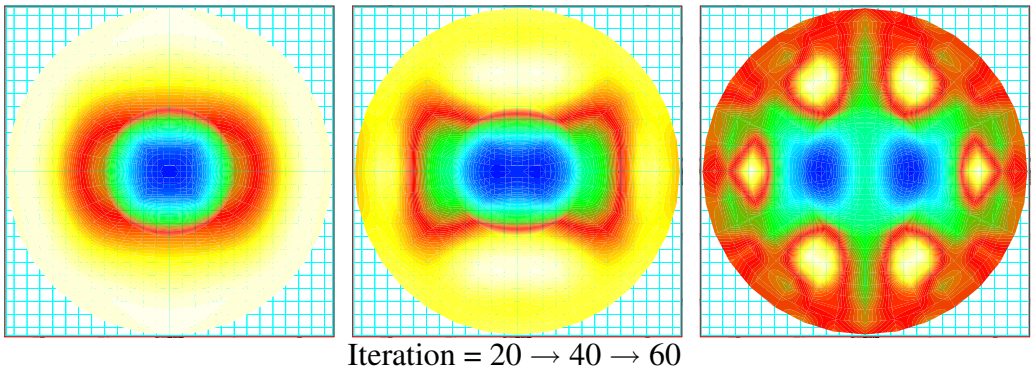


Figure 7.13: A two dimensional structured finite element circular mesh of diameter 5cm with two irregular shaped void regions embedded deeply in the domain. The steady state problem has 1186 nodes and 1760 bi-linear quadrilateral elements.

Eight equally spaced isotropic sources and detectors are situated around the circumference of the medium to obtain the forward model and reconstruction model results. A P_3 angular approximation is used to obtain the steady state forward and inversion calculations along with a Levenberg-Marquardt optimisation method. The finite element mesh used for the forward and inverse reconstruction contained 1186 nodes and 1760 bi-linear quadrilateral elements. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set 10,000 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. The x co-ordinate and y co-ordinate positions of each of the sources and detectors placed around the circumference of the medium are illustrated in table 7.1.

Figure 7.14 compares the scattering reconstruction results and figure 7.15 compares the absorption reconstruction results for the P_3 angular approximation after 20 iterations, 40 iterations, 60 iterations, 80 iterations and 100 iterations respectively. Figures 7.14 and 7.15 provide visual evidence that during P_3 inversion calculations, the algorithm finds the correct location and approximate size of the irregular shaped voids for both the absorption



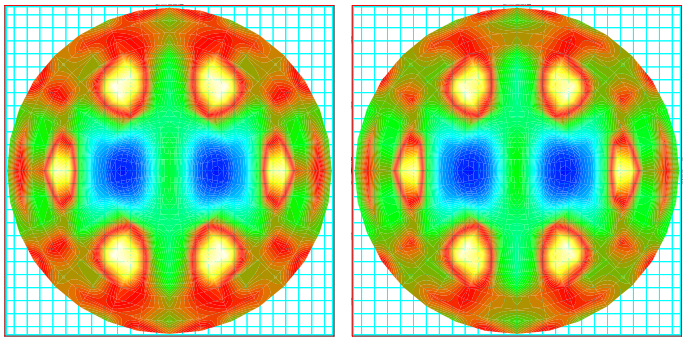
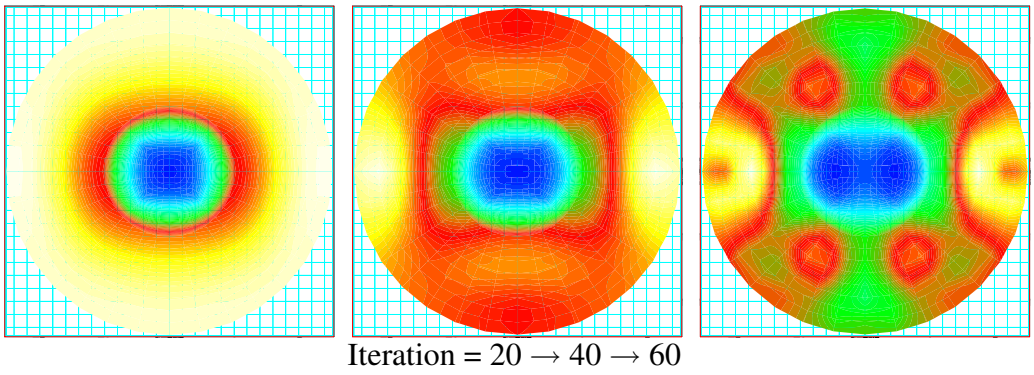


Figure 7.14: Scattering reconstruction results for two irregular shaped voids embedded inside a circle of diameter 5cm. A P_3 angular expansion is used to perform the image reconstruction. The plots illustrate material property reconstructions after 20, 40, 60, 80 and 100 iterations.

and scattering reconstructions.

The relative magnitudes of the absorption and scattering coefficients are approximated fairly accurately but not as accurately as the absorption and scattering coefficients are resolved in figure 7.10. This is largely due to the increased complexity of the problem. Figure 7.16 illustrates the absorption (top image) and scattering (bottom images) images after 100 iterations for the reconstructions described in figures 7.14 and 7.15. The results approximate the magnitude of the absorption coefficient, μ_a, to extend from 0.272cm^{-1} to approximately 0.116cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 13.8cm^{-1} to approximately 1.48cm^{-1}. The magnitudes of both the scattering and absorption coefficients of the inhomogeneities indicate that a void is present but are overestimated. The magnitude of the background scattering coefficient is overes-



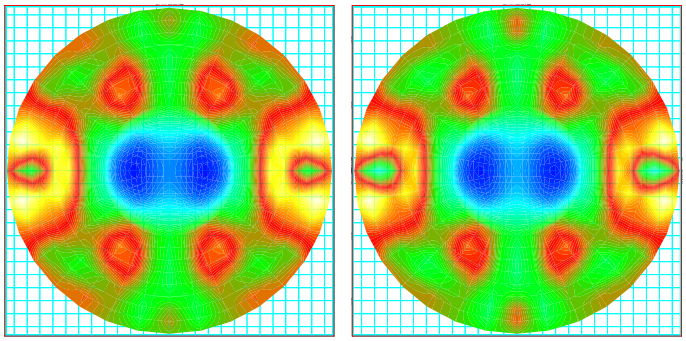


Figure 7.15: Absorption reconstruction results for two irregular shaped voids embedded inside a circle of diameter 5cm. A P_3 angular expansion is used to perform the image reconstruction. The plots illustrate material property reconstructions after 20, 40, 60, 80 and 100 iterations.

timated in regions where the six image artefacts exist, otherwise the background medium is well approximated. The image artefacts occur in the vicinity of the void regions.

Figure 7.17 illustrates the error functional as a function of iteration number for the first 100 iterations of the P_3 angular approximation described in figure 7.13 and table 7.8. The error functional is reduced by almost three orders of magnitude during the 100 iteration calculation. The error functional is reduced gradually from a magnitude of $F = 1.84$ to a magnitude of $F = 0.0064$. No sharp drop in magnitude is observed. The gradual reduction in the error functional is because of the high initial value of the regularisation penalty levels, γ_a and γ_{s0}. The regularisation penalty levels are initially set to 10,000 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. A positive aspect of reducing the regularisation penalty levels by a larger annealing factor

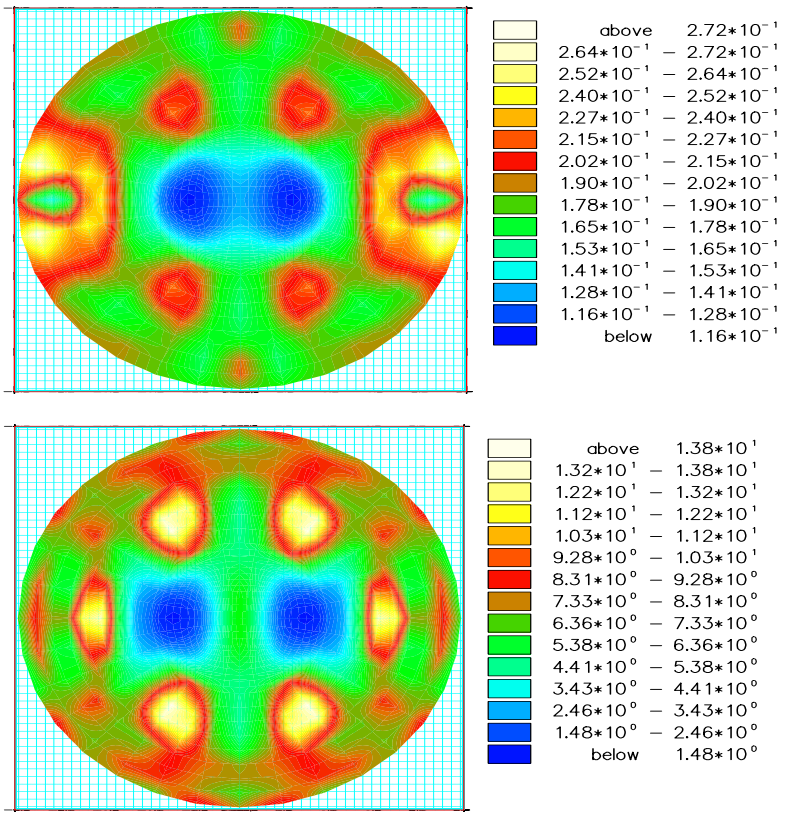


Figure 7.16: Absorption (top image) and scattering (bottom image) reconstruction results for a circular problem containing two irregular shaped void regions. A P_3 angular approximation is used. The inversion was performed for 100 iterations using eight equidistantly spaced isotropic sources and detectors.

than the one used in the above example is that material structure is likely to be resolved earlier in the inversion. However, more noise and a greater degree of ill-posedness could be introduced into the imaging domain and could result in a greater number of image artefacts or more predominant image artefacts.

Transport theory scattering and absorption image reconstructions may be improved by introducing extra information into the imaging domain. Extra information can be introduced into the imaging domain by increasing the number of sources and detectors placed around the circumference of the medium. Figure 7.18 illustrates the absorption (top image) and scattering (bottom image) reconstruction images for the cylindrical problem described in figure 7.13 and table 7.8. A P_3 angular approximation is used to obtain the

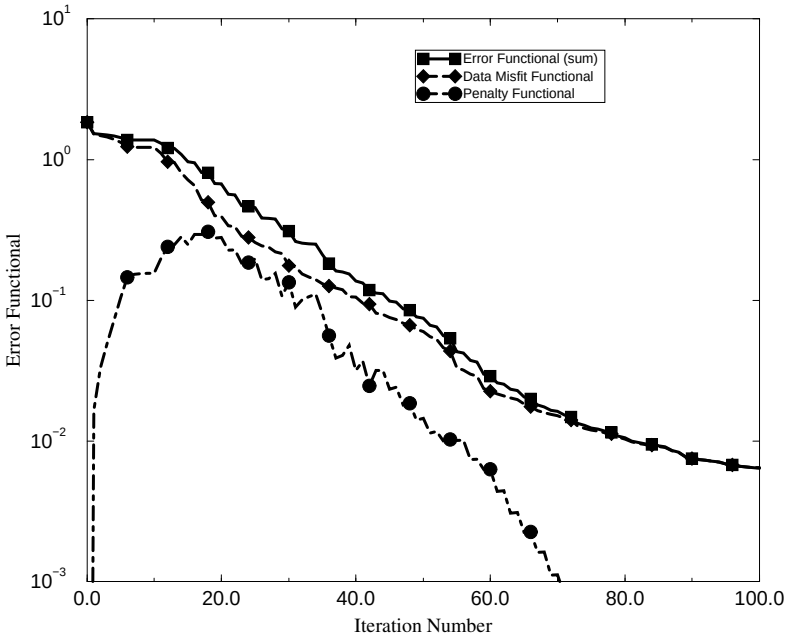


Figure 7.17: Error functional as a function of iteration number for the first 100 iterations of the P_3 angular approximation described in figure 7.13 and table 7.8.

steady state forward and inversion model results along with a Levenberg-Marquardt optimisation method. The inversion was performed for 100 iterations using twelve equally spaced isotropic sources and detectors situated around the circumference of the medium. All other parameters are the same as those previously stated in the experiment performed using eight sources and eight detectors.

The algorithm finds the correct location and approximate size of the the irregular shaped voids for both the absorption and scattering reconstructions. The relative magnitudes of the absorption and scattering coefficients are also approximated fairly accurately. The results approximate the magnitude of the absorption coefficient, μ_a, to extend from 0.243cm^{-1} to approximately 0.109cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 13.3cm^{-1} to approximately 2.22cm^{-1}. The magnitudes of both the scattering and absorption coefficients of the inhomogeneities indicate that a void is present but is overestimated. While the relative absorption and scattering magnitudes are similar to those estimated for the problem geometry that uses eight sources and detectors,

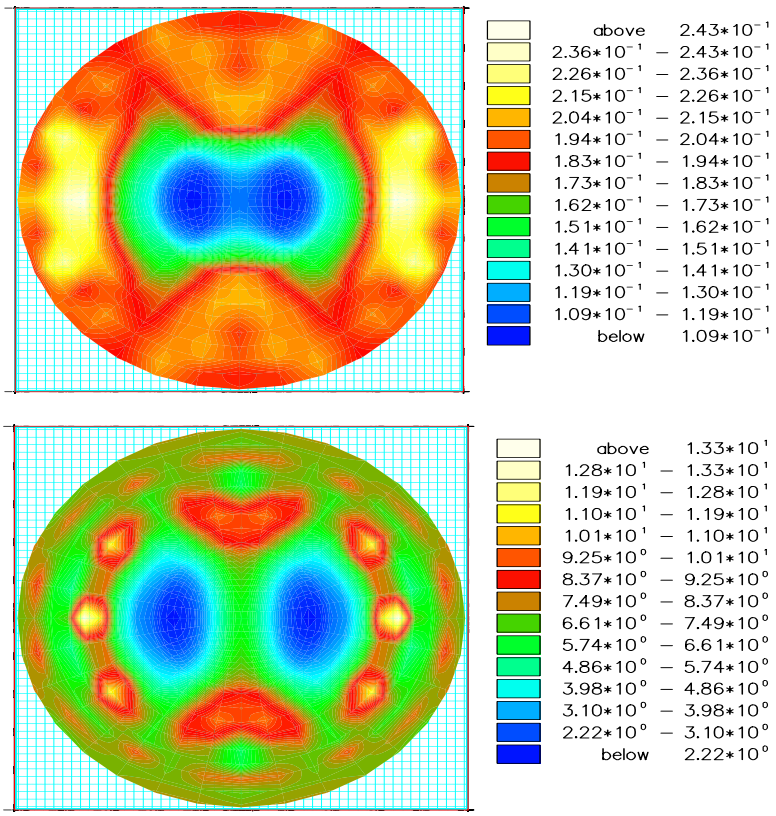


Figure 7.18: Absorption (top image) and scattering (bottom image) reconstruction results for a circular problem containing two irregular shaped transparent regions. A P_3 angular approximation is used. The inversion was performed for 100 iterations using twelve equidistantly spaced isotropic sources and detectors.

the image aftefacts are reduced. A slight improvement in the image quality is therefore observed.

Figure 7.19 depicts absorption coefficient (top graph) and scattering coefficient (bottom graph) as a function of distance along the y-axis for the P_3 angular approximation described in figures 7.14, 7.15, and 7.18. A line is drawn along the y-axis, at co-ordinates (0.0, -2.5), (0.0, 2.5). The solid line represent results obtained for calculations performed using 12 sources and 12 detectors. The dashed line represent results obtained for calculations performed using 8 sources and 8 detectors and the dotted line represents the exact solution. The inversion was performed for 100 iterations. In agreement with figures 7.14, 7.15, and 7.18 the two irregular void regions are clearly resolved for both absorption and

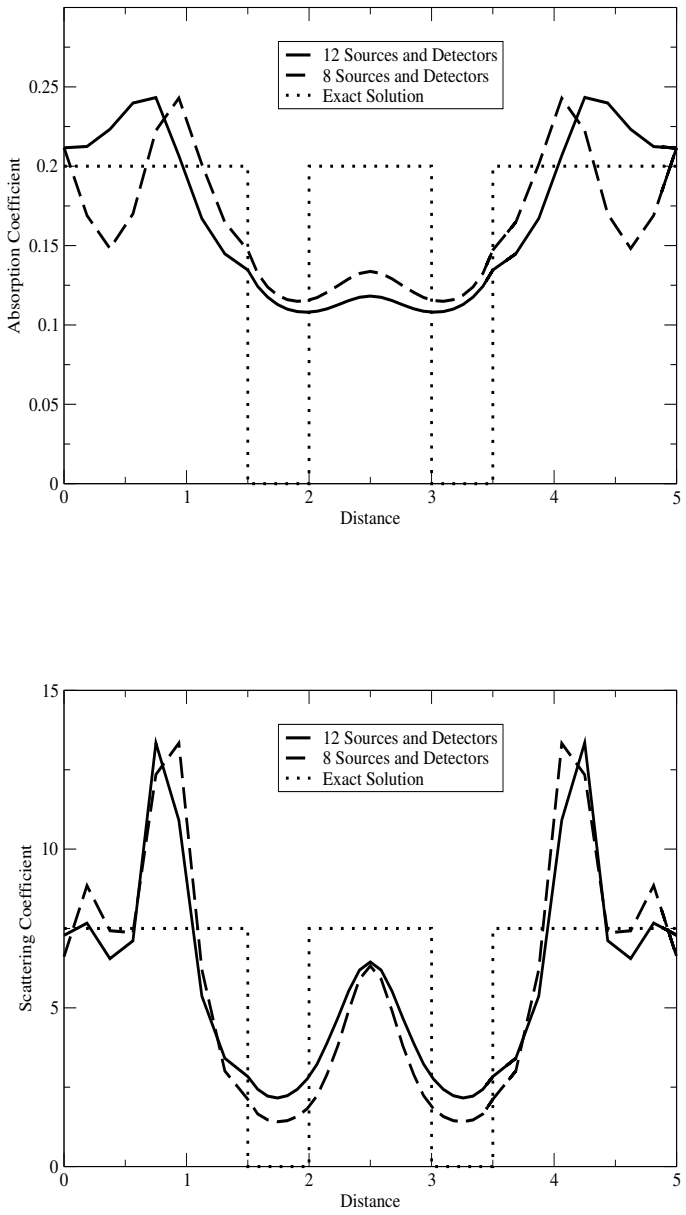


Figure 7.19: Absorption coefficient (top graph) and scattering coefficient (bottom graph) as a function of distance along the y-axis for the P_3 angular approximation described in figures 7.14, 7.15 and 7.18. A line is drawn along the y-axis, at co-ordinates (0.0, -2.5), (0.0, 2.5). The solid line represent results obtained for calculations performed using 12 sources and 12 detectors. The dashed line represent results for obtained for calculations performed using 8 sources and 8 detectors and the dotted line represents the exact solution. The inversion was performed for 100 iterations.

scattering reconstructions. While the voids associated with scattering reconstructions are the correct approximate size, the voids associated with absorption reconstructions provide less accurate results. The material located between the void regions is also resolved more accurately for scattering reconstructions than for absorption reconstructions. Furthermore, the graph also cuts through the region where two of the six image artefacts exist. These are illustrated in the peaks that extend beyond magnitudes of 12.5cm^{-1} for scattering reconstructions and 0.272cm^{-1} for absorption reconstructions.

Problem 3

The simulation geometry described in figure 7.20 depicts a circle 7cm in diameter containing an embedded void-like ring of thickness 0.5cm. The ring extends from a radius of 2cm to a radius of 2.5cm from the centre of the domain. The absorption and scattering coefficients of the background medium and the inhomogeneous ring are contained in table 7.9. The problem geometry represents a crude approximation of the neonatal brain. The optical absorption and scattering coefficients of the background medium approximate those of the brain of a 24 week gestation neonate [133], while a cylindrical diameter of 7cm represents a preterm neonate's approximate head size. The transparent ring embedded inside the neonatal brain represents the cerebrospinal fluid that surrounds the human brain. It's relative thickness and location are arbitrary. Eight equally spaced isotropic sources and detectors are placed around the circumference of the medium to obtain the forward model and reconstruction model results. P_1 and P_3 angular approximations are used to obtain the steady state forward and inversion calculations. The same sized mesh is used to generate both the synthetic data and the inversion reconstruction. The mesh used for the forward and inverse reconstruction contained 2210 nodes and 3296 bi-linear quadrilateral elements. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 100 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. The x co-ordinate and y co-ordinate positions of each of the sources and detectors placed around the circumference of the medium are illustrated in table 7.1.

Figure 7.21 compares the converged absorption (left hand side) and scattering (right hand

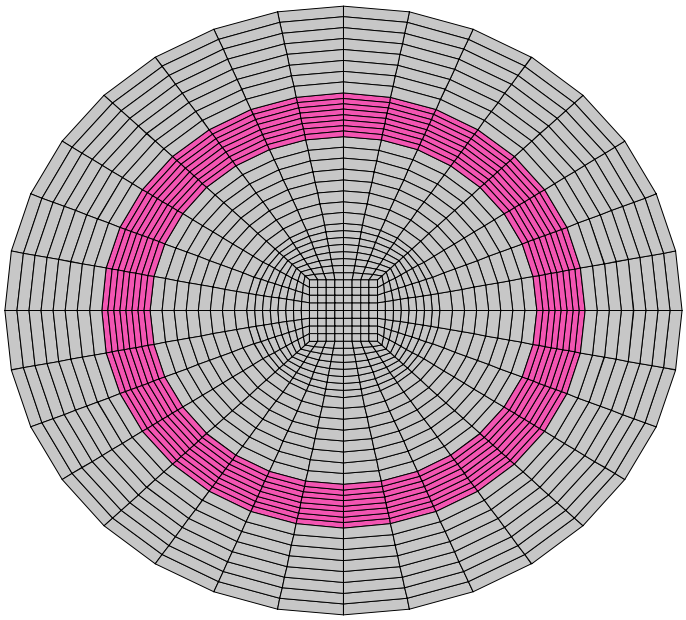


Figure 7.20: A two dimensional structured finite element circular mesh of diameter 7cm with an embedded void-like ring of thickness 0.5cm. The ring extends between a radius of 2cm to a radius of 2.5cm from the centre of the domain. The steady state problem has 2210 nodes and 3296 bi-linear quadrilateral elements.

	$\mu_a[cm^{-1}]$	$\mu_s^{'}[cm^{-1}]$
Background Medium	0.2	7.5
Ring Region	0.0	0.0

Table 7.9: Optical absorption and scattering coefficients for a circle 7cm in diameter containing an embedded void-like ring of thickness 0.5cm.

side) reconstruction results for the problem geometry described in figure 7.20 and table 7.9. P_1 (top) and P_3 (bottom) angular approximations were used to perform the inversion calculations. The results confirm that for steady state void like configurations -in which diffusion theory is not valid- full transport theory is necessary to reconstruct an objects material structure. A P_3 angular expansion or above approximates full transport theory. Figure 7.21 provides visual evidence that during the P_3 scattering reconstruction, the algorithm finds the correct location of the transparent ring. The thickness of the void is overestimated. The P_3 angular approximation fails to reconstruct the absorption material structure. Instead of reconstructing a ring-like region, a large circular void is reconstructed similar to the circular void described in figures 7.9 and 7.10. The P_1 angular approximation fails to reconstruct both the scattering and absorption material structure. According to the P_1 images described in figure 7.21, the ring-like structure described in the original problem geometry (figure 7.20) does not exist. The results suggest that the diffusion approximation to the steady state Boltzmann transport equation has limited use when considering void modelling. The diffusion approximation is therefore largely inappropriate for modelling neonatal head phantoms.

The relative magnitudes of the absorption and scattering coefficients are approximated fairly accurately although they are never completely resolved. The P_1 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.182cm^{-1} to approximately 0.0147cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.41cm^{-1} to approximately 0.0871cm^{-1}. The P_3 angular expansion approximates the magnitude of the absorption coefficient, μ_a, to extend from 0.214cm^{-1} to approximately 0.0232cm^{-1}, while the magnitude of the scattering coefficient extends from approximately 8.41cm^{-1} to approximately 0.527cm^{-1}. The legend illustrated at the bottom of figure 7.21 depicts the range in magnitude of the absorption coefficient (left hand side) and scattering coefficient (right hand side) for the P_3 angular approximation described above. For both P_1 and P_3 angular approximations, the magnitudes of both the scattering and absorption coefficients of the inhomogeneous ring indicate that a void is present but is overestimated. The magnitude of the background scattering coefficient is also overestimated for both P_1 and P_3 angular approximations. The magnitude of the scattering coefficient reconstruction of the internal cylinder (contained on the inside of

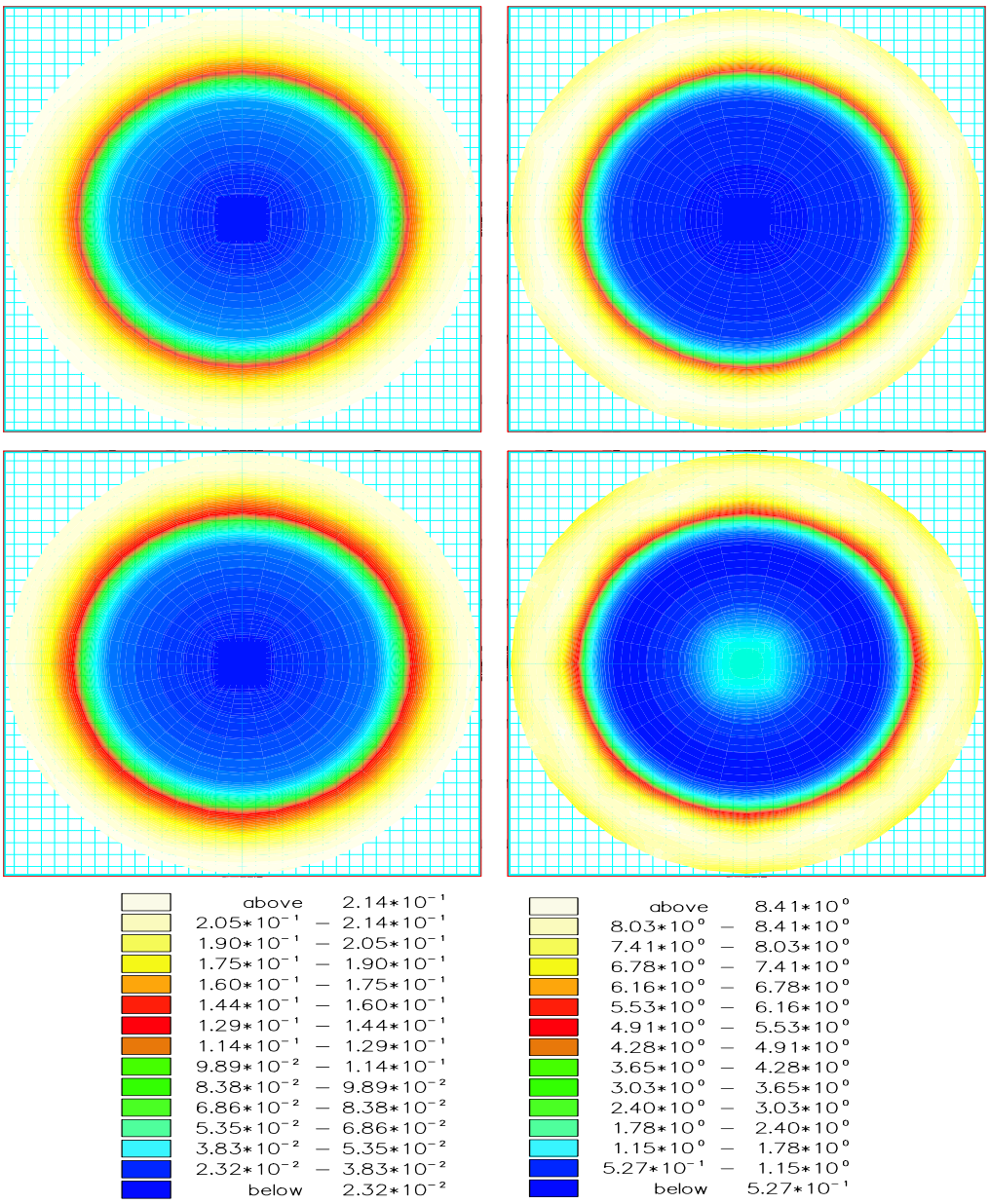


Figure 7.21: Absorption (left hand side) and scattering (right hand side) reconstruction results for a circular problem containing an embedded transparent ring. P_1 (top) and P_3 (bottom) angular expansions are used to obtain the steady state reconstruction results. The calculations took 13 iterations and 24 iterations to converge during the P_1 and P_3 angular approximations respectively. The legend illustrated at the bottom of the figure depicts the range in magnitude of the absorption coefficient (left hand side) and scattering coefficient (right hand side) for the P_3 angular approximation.

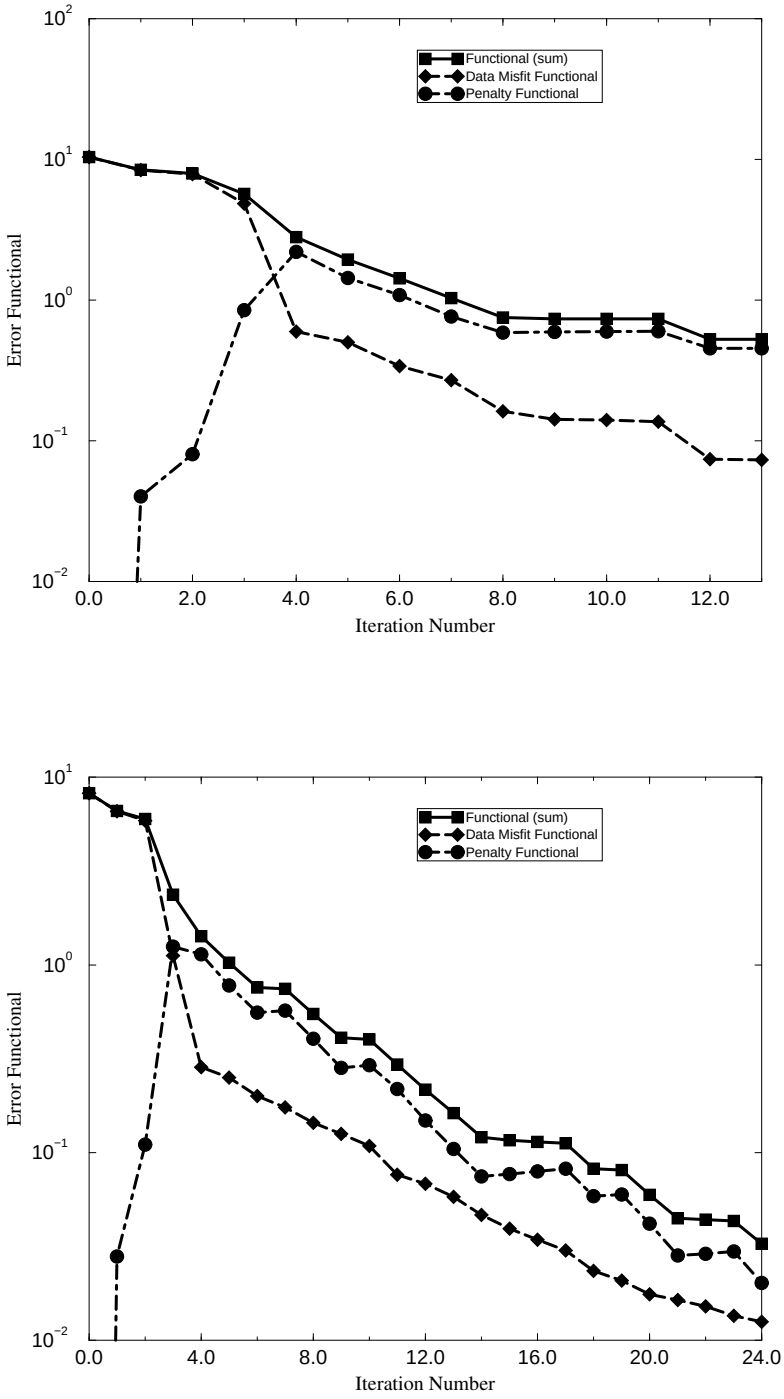


Figure 7.22: Error functional as a function of iteration number. During both P_1 (top graph) and P_3 (bottom graph) angular approximations the error functional is reduced most dramatically during early stages of the inversion. A sharper descent in the error functional is observed for the P_3 angular approximation.

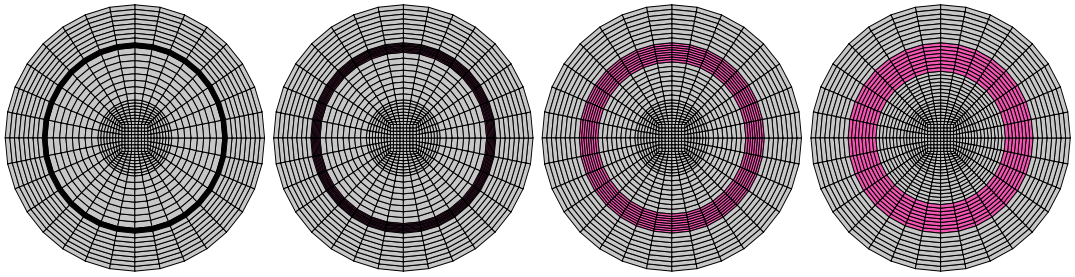
the inhomogeneous ring), described in the P_3 reconstruction is slightly underestimated.

Transport theory scattering and absorption image reconstructions may be improved by using a higher order angular approximation (i.e. P_5, P_7, P_9, etc ....) or by introducing more information into the imaging domain. Additional information can be introduced into the imaging domain by using more sources and detectors, incorporating time-dependent information, applying a sophisticated regularisation scheme or by developing a tracer distribution/fluorescence method.

Figure 7.22 illustrates the error functional as a function of the iteration number using P_1 (top graph) and P_3 (bottom graph) angular approximations. The P_1 angular approximation converged after 13 iterations from an error functional of $F = 10.42$ to an error functional of $F = 0.0733$. The P_3 angular approximation converged after 24 iterations from an error functional of $F = 8.22$ to an error functional of $F = 0.012$. The convergence criterion is chosen so that once convergence is met, increasing the number of iterations does not change the inversion model significantly. In other words, even after 50 iterations the ring-like material structure is never resolved for P_1 calculations. For both P_1 and P_3 angular approximations the error functional is reduced most dramatically during early stages of the inversion (i.e. during the first six iterations). The penalty contribution to the error functional takes greatest effect after the third iteration. Deeply embedded material structure emerges when the penalty contribution to the error functional is larger than the data misfit contribution to the error functional.

Incorporating Time Dependent Information into the Imaging Domain

Using a time dependent signal provides a means of introducing more information into the imaging domain [44]. Transient inversion may therefore be used to improve the quality of absorption and scattering coefficient reconstructions inside a given medium. The simulation geometry described in figure 7.23 depicts a series of circular problems, 7cm in diameter. The circular domains contain embedded transparent rings of thicknesses (from left to right) 0.125cm, 0.25cm, 0.5cm (like in the previous steady state example) and 0.75cm respectively. The rings are embedded 1cm deep into the respective media. The absorption and scattering coefficients of the background medium and each of the embedded voids are the same as those described in table 7.9. For each of the model geometries,

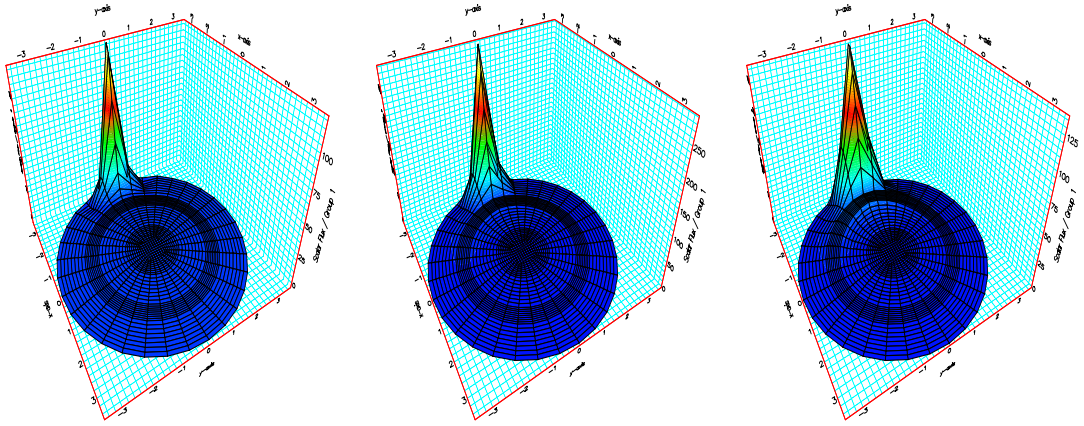


Void Thickness (from left to right) - 0.125cm → 0.25cm → 0.5cm → 0.75cm

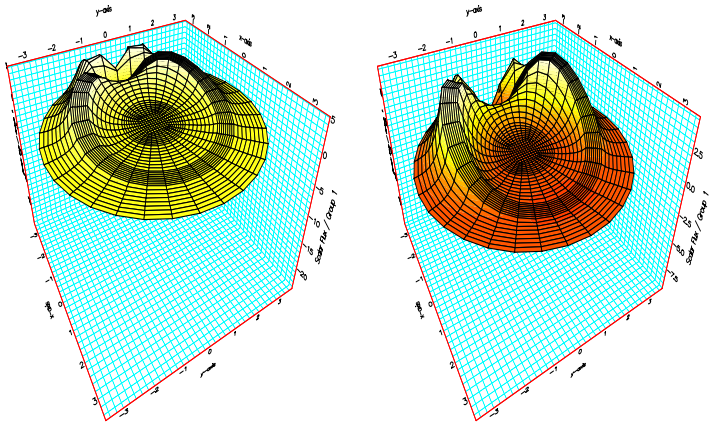
Figure 7.23: Two dimensional structured finite element circular meshes of diameter 7cm with embedded void rings of thicknesses (from left to right) 0.125cm, 0.25cm, 0.5cm, and 0.75cm. The rings are embedded 1.0cm from the surface of the domain. Each of the time dependent problems have 2210 nodes, 3296 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps.

the photon velocity is set arbitrarily to 1 and a time step of $\Delta t = 5$ is used. Eight equally spaced isotropic sources and detectors are placed around the circumference of the medium to obtain the forward model and inverse model reconstruction results. The sources and detectors are located in the same positions as those used to obtain the steady state reconstruction results described in figure 7.21. A P_1 angular approximation is used to obtain the time dependent forward model and inversion model calculations along with a Levenberg-Marquardt optimisation method. The finite element mesh used has 2210 nodes, 3296 bi-linear quadrilateral elements and 5 Discontinuous Galerkin time steps. Both the absorption and scattering regularisation penalty levels (γ_a and γ_{s0}) are initially set to 10,000 and are annealed downwards by a factor of 1.5 after each Levenberg-Marquardt iteration. The step length damping coefficient, λ, is initially set to 0.1. As an initial guess, a homogeneous medium is chosen with optical properties that equal the background. In other words $\mu_a = 0.2$cm and $\mu_s^{'} = 7.5$cm.

A Discontinuous Galerkin technique is used to discretise the time domain. The Discontinuous Galerkin method has the ability to use large time steps while still capturing the essential features of the time dependent signal [44]. Figure 7.24 depicts scalar flux as a function of two dimensional space (x and y) for the exact solution at time levels 1 (top left hand side graph), 2 (top middle graph), 3 (top right hand side graph), 4 (bottom left



Time Levels 1 → 2 → 3



Time Levels 4 → 5

Figure 7.24: Scalar flux as a function of two dimensional space (x and y) for the exact solution at time levels 1 (top left hand side graph), 2 (top middle graph), 3 (top right hand side graph), 4 (bottom left hand side graph), and 5 (bottom right hand side graph) respectively. A P_1 angular approximation has been used to obtain each of the above images. The scalar flux extends up to a maximum magnitude of over $100cm^{-2}s^{-1}$ during the first time step, $250cm^{-2}s^{-1}$ during the second time step, $125cm^{-2}s^{-1}$ during the third time step, $5cm^{-2}s^{-1}$ during the fourth time step and $25cm^{-2}s^{-1}$ during the fifth time step.

hand side graph), and 5 (bottom right hand side graph) respectively. A single source is illustrated and is placed at x- y- co-ordinate position (-3.5, 0.0). The transparent ring has a thickness of 0.5cm. The numerical delta function signal begins as a large, slim peak before propagating through the medium as time increases. Figure 7.24 illustrates how the scalar flux disperses through the medium and how the magnitude of the peak scalar flux decreases as a function of time. The scalar flux extends up to a maximum magnitude of over $100cm^{-2}s^{-1}$ during the first time step, $250cm^{-2}s^{-1}$ during the second time step, $125cm^{-2}s^{-1}$ during the third time step, $5cm^{-2}s^{-1}$ during the fourth time step and $25cm^{-2}s^{-1}$ during the fifth time step. A noteworthy feature is the channeling of the scalar flux through the transparent ring during time levels 4 and 5.

Figure 7.25 compares the absorption (left hand side images) and scattering (right hand side images) reconstruction results for each of the problem geometries described in figure 7.23. A P_1 angular approximation was used to perform the transient forward model and inversion model calculations. The reconstruction was performed for 50 iterations. The top set of images describe the absorption and scattering reconstruction results for a ring thickness of 0.125cm. The top middle set of images describe the absorption and scattering reconstruction results for a ring thickness of 0.25cm. The bottom middle set of images describe the absorption and scattering reconstruction results for a ring thickness of 0.5cm, while the bottom set of images describe the absorption and scattering reconstruction results for a ring thickness of 0.75cm. Figure 7.25 provides visual evidence that during the P_1 inversion calculation, the algorithm finds the correct location and approximate size of the transparent ring for both absorption and scattering reconstructions providing the transparent ring does not exceed a thickness of 0.5cm. The most accurate reconstruction images are observed for ring thicknesses of 0.25cm and 0.5cm.

Figure 7.25 illustrates that incorporating time dependent information into the imaging domain has a significant effect on image quality. When comparing reconstructions performed using steady state radiation transport theory (figure 7.21) to reconstructions performed using transient radiation transport theory (figure 7.25) it is clear that visualisations obtained using time dependent information appear sharper and more accurate than those obtained using steady state transport. The scattering coefficient reconstruction of the transparent ring appears clearer and the thickness appears more accurately resolved when

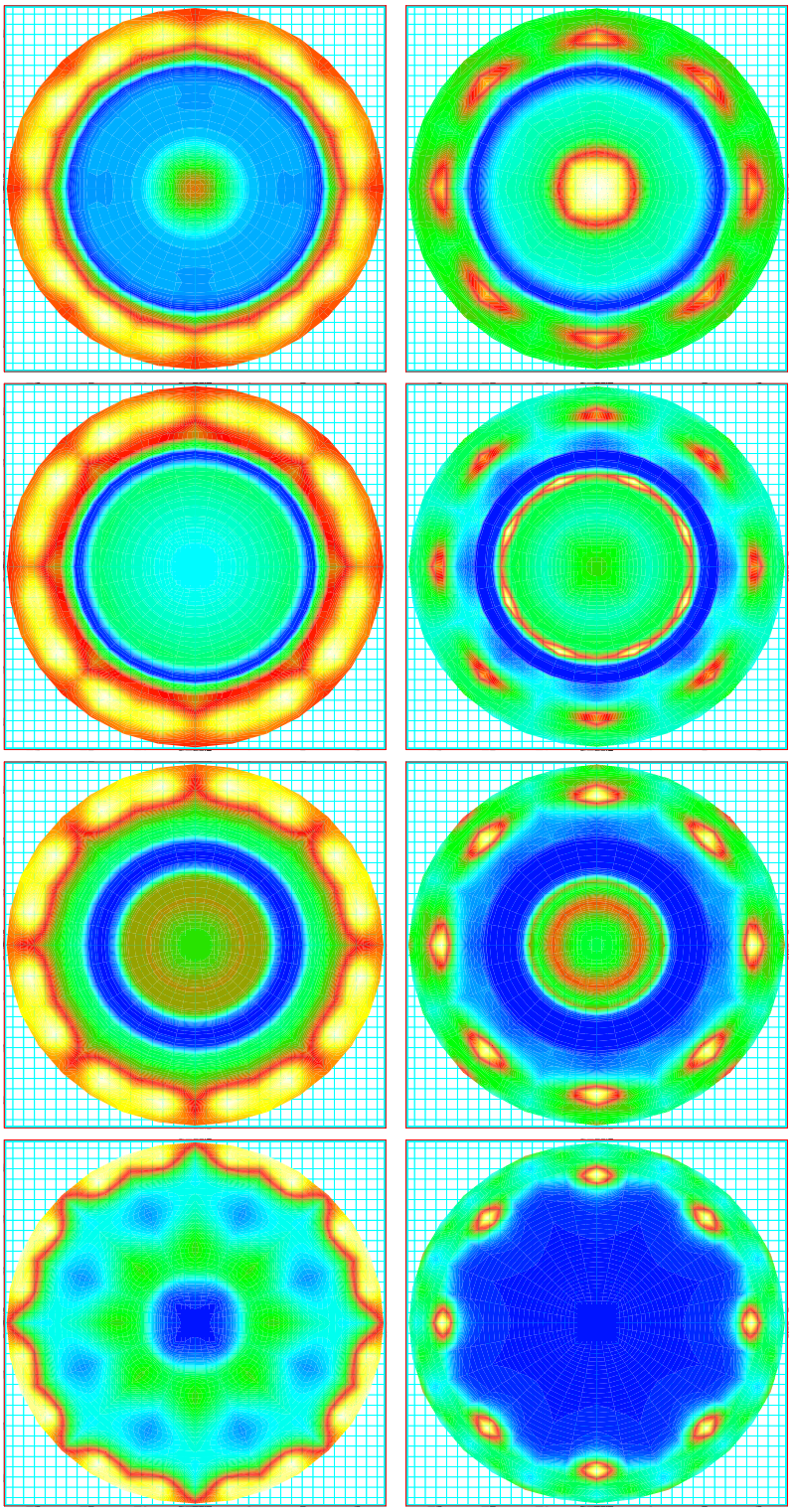


Figure 7.25: Absorption (left hand side) and scattering (right hand side) reconstruction results for a circular problem containing an embedded void ring. Ring thicknesses of 0.125cm (top images), 0.25cm (top middle image), 0.5cm (bottom middle image) and 0.75cm (bottom image) are investigated. A time-dependent P_1 angular expansion is performed for 50 iterations.

comparing P_1 transient information to P_3 steady state information. During steady state radiation transport the absorption coefficient reconstruction fails to reconstruct a ring-like structure during both P_1 and P_3 angular expansions.

Figures 7.26 and 7.27 depict absorption coefficient (left hand graphs) and scattering coefficient (right hand graphs) reconstruction results as a function of distance along the x-axis for the P_1 angular approximations described in figure 7.25. Figure 7.26 illustrates the inversion model results for ring thicknesses of 0.125cm (top set of graphs) and 0.25cm (bottom set of graphs), while figure 7.27 illustrates the inversion model results for ring thicknesses of 0.5cm (top set of graphs) and 0.75cm (bottom set of graphs). During each of the image reconstructions, a line is drawn from the middle of the y-axis, at co-ordinates (0.0, -2.5), (0.0, 2.5). The solid line represents the exact solution, while the dashed line represents the reconstruction result after 50 iterations.

Figures 7.26 and 7.27 confirm that the most accurate image reconstruction results are obtained for void thicknesses of 0.25cm and 0.5cm. The relative magnitudes of the absorption and scattering coefficients are approximated fairly accurately. For ring thicknesses of 0.25cm and 0.5cm, the magnitude of the absorption coefficient, μ_a, ranges from approximately 0.15cm^{-1} to approximately 0.02cm^{-1}, while the magnitude of the scattering coefficient ranges from approximately 15.0cm^{-1} to approximately 0.15cm^{-1}. The magnitudes of both the absorption and scattering coefficient reconstructions indicate that a void is present but are both overestimated inside the transparent void. Another noteworthy observation is that the thickness of the transparent void is most accurately resolved during absorption coefficient reconstructions. During scattering coefficient reconstructions the thickness of the transparent void is over approximated.

The top set of graphs illustrated in figure 7.26 describe the reconstruction results for a transparent void of thickness 0.125cm. While a ring of the correct approximate thickness is reconstructed for both absorption and scattering coefficient reconstructions the results look less visually pleasing than those reconstructed for rings of thickness 0.25cm and 0.5cm respectively. The relative magnitudes of the absorption and scattering coefficients are approximated fairly accurately . The magnitude of the absorption coefficient, μ_a, ranges from approximately 0.15cm^{-1} to approximately 0.05cm^{-1}, while the magni-

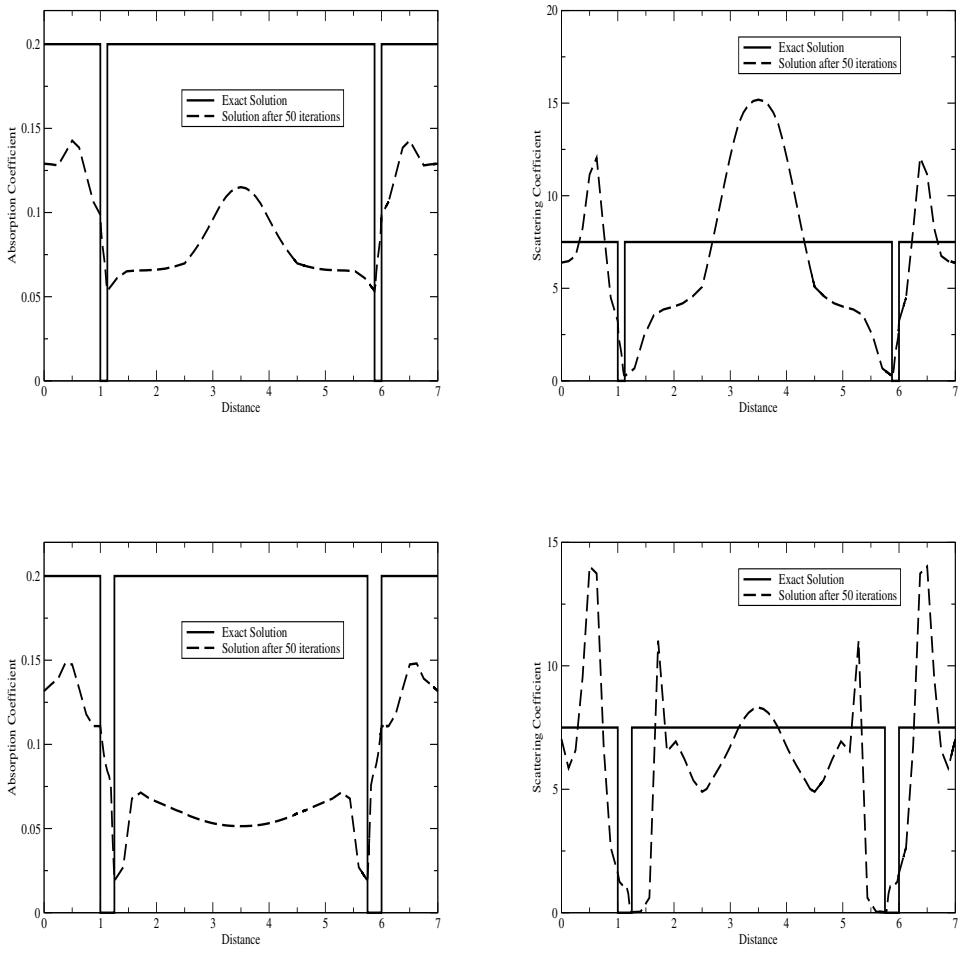


Figure 7.26: Absorption coefficient (left hand side) and scattering coefficient (right hand side) as a function of distance along the x-axis for a transparent ring of thickness 0.125cm (top set of graphs) and a transparent ring of thickness 0.25cm (bottom set of graphs). The solid lines represent the exact solution, while the dashed line represents the transient P_1 reconstruction after 50 iterations.

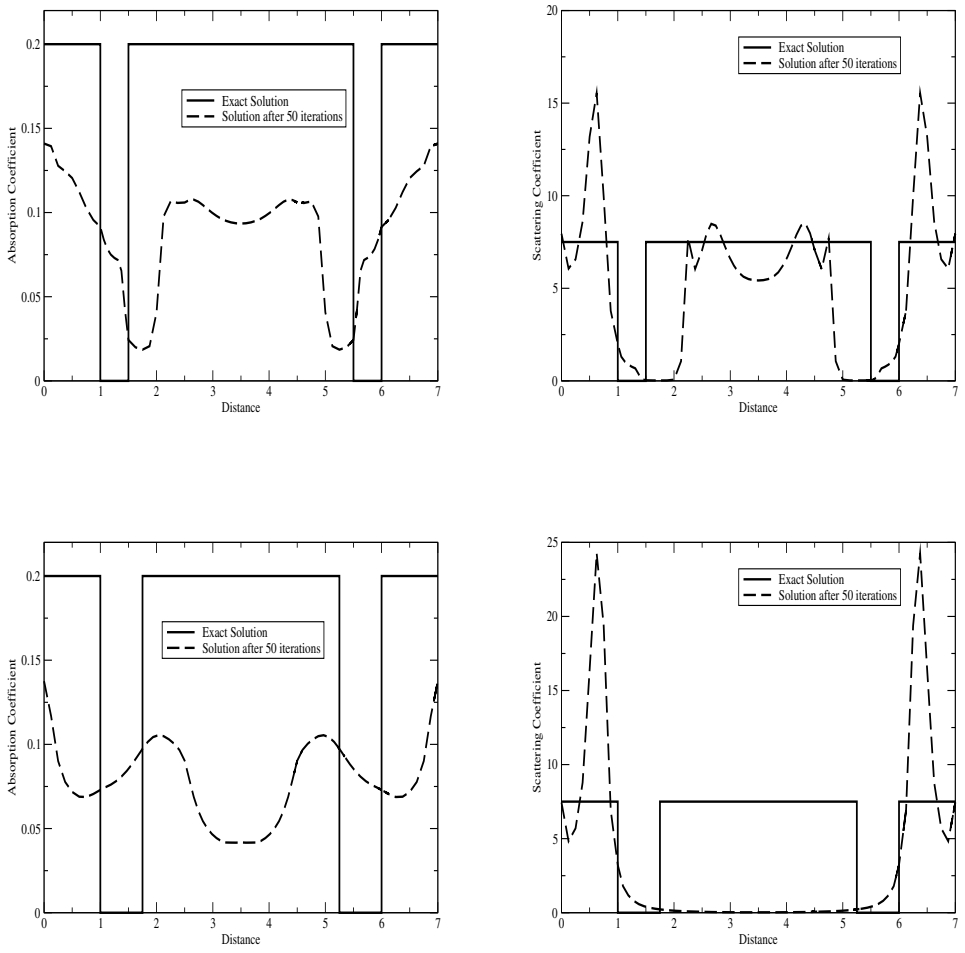


Figure 7.27: Absorption coefficient (left hand side) and scattering coefficient (right hand side) as a function of distance along the x-axis for a transparent ring of thickness 0.5cm (top set of graphs) and a transparent ring of thickness 0.75cm (bottom set of graphs). The solid lines represent the exact solution, while the dashed line represents the transient P_1 reconstruction after 50 iterations.

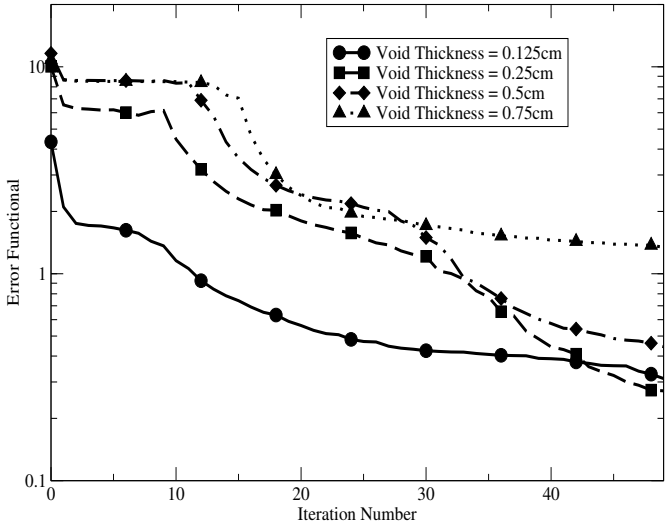


Figure 7.28: Reduction of the error functional during the first 50 iterations of the transient inversion illustrated in figure 7.23. The solid line represents the reduction in error functional for a transparent ring of thickness 0.125cm, the dashed line represents the reduction in error functional for a transparent ring of thickness 0.25cm, the dot-dashed line represents the reduction in error functional for a transparent ring of thickness 0.5cm and the dotted line represents the reduction in error functional for a transparent ring of thickness 0.75cm.

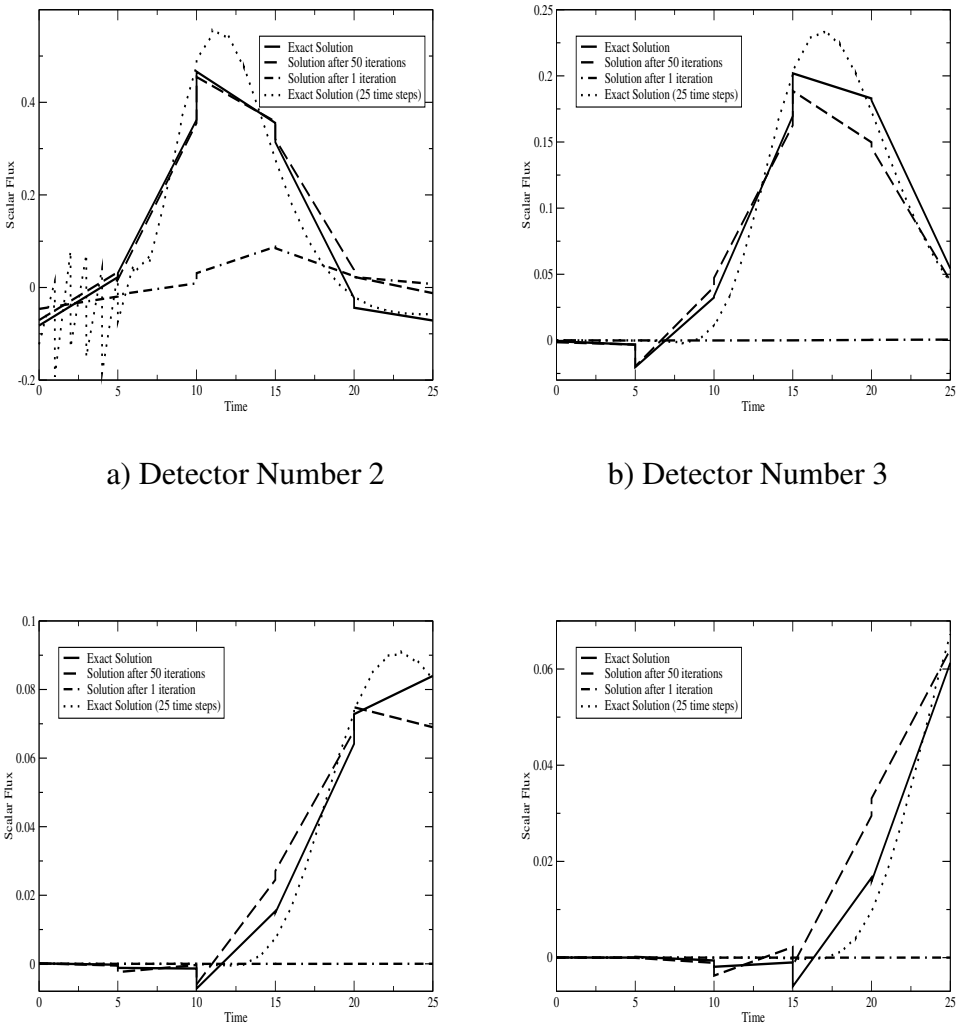
tude of the scattering coefficient ranges from approximately 15.0cm^{-1} to approximately 0.11cm^{-1}. Inversion calculations performed on a void of thickness 0.75cm fail to reconstruct a useful image during both absorption and scattering reconstructions.

During each of the image reconstructions described in figures 7.25, 7.26 and 7.27 the magnitude of the background scattering coefficient is overestimated (by varying degrees) in regions where eight image artefacts exist. The eight image artefacts are located in the vicinity of where the eight sources and eight detectors are situated. The scattering coefficient magnitudes contained in these regions could be improved either by increasing the regularisation penalty levels, γ_a and γ_{s0}, or by utilising a more sophisticated regularisation strategy.

Figure 7.28 illustrates error functional as a function of the iteration number during the first 50 iterations of the time dependent P_1 angular approximations described above. The solid line represents the reduction in error functional for a transparent ring of thickness 0.125cm, the dashed line represents the reduction in error functional for a transparent ring of thickness 0.25cm, the dot-dashed line represents the reduction in error functional for a transparent ring of thickness 0.5cm and the dotted line represents the reduction in error functional for a transparent ring of thickness 0.75cm. The error functional is reduced by two orders of magnitude for void thicknesses of 0.25cm and 0.5cm and by one order of magnitude for void thicknesses of 0.125cm and 0.75cm. The reduction of the error functional therefore supports the observation that void thicknesses of 0.25cm and 0.5cm provide better image reconstruction results.

Figure 7.29 illustrates scalar flux as a function of time for the first source at the second detector location (top left hand graph), the third detector location (top right hand graph), the fourth detector location (bottom left hand graph) and fifth detector location (bottom right hand graph). Figure 7.30 illustrates scalar flux as a function of time for the first source at the sixth detector location (top left hand graph), the seventh detector location (top right hand graph), and the eighth detector location (bottom left hand graph). The eight isotropic sources and detectors are placed equidistantly around the circumference of the domain. The exact scalar flux solution obtained using 5 Discontinuous Galerkin time steps (solid line) is compared alongside the exact scalar flux solution obtained using 25 Discontinuous Galerkin time steps (dotted line) to illustrates that the Discontinuous Galerkin method has the ability to use large time steps while still capturing the essential features of the time dependent signal. The co-ordinate positions of each of the individual sources and detectors are described in table 7.1. Figures 7.29 and 7.30 compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 50 iterations (dashed line). The detector response received at the first detector has been omitted from the calculations (using the weight function) because the large source-detector sensitivities observed when sources and detectors are placed too close together are not regarded as reliable. The first detector has been placed on top of the first source.

Important features illustrated in figures 7.29 and 7.30 include the symmetry of the mod-



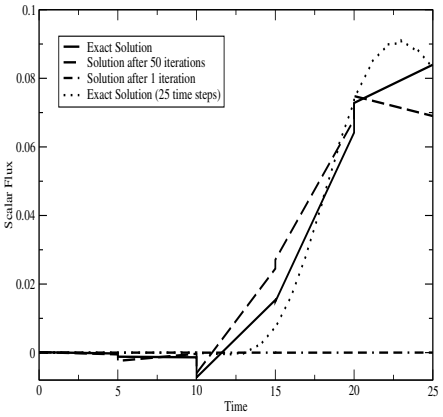
a) Detector Number 2

b) Detector Number 3

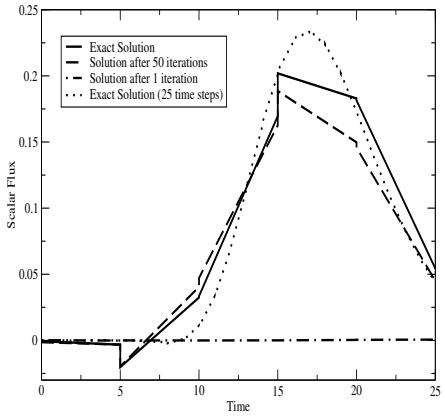
c) Detector Number 4

d) Detector Number 5

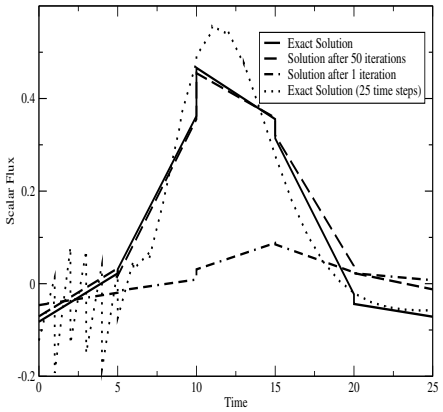
Figure 7.29: Scalar flux versus time for the first source at a) the second detector location, b) the third detector location, c) the fourth detector location and d) the fifth detector location. The graphs compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 50 iterations (dashed line). Calculations were performed for a void of thickness 0.25cm. The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each graph.



a) Detector Number 6



b) Detector Number 7

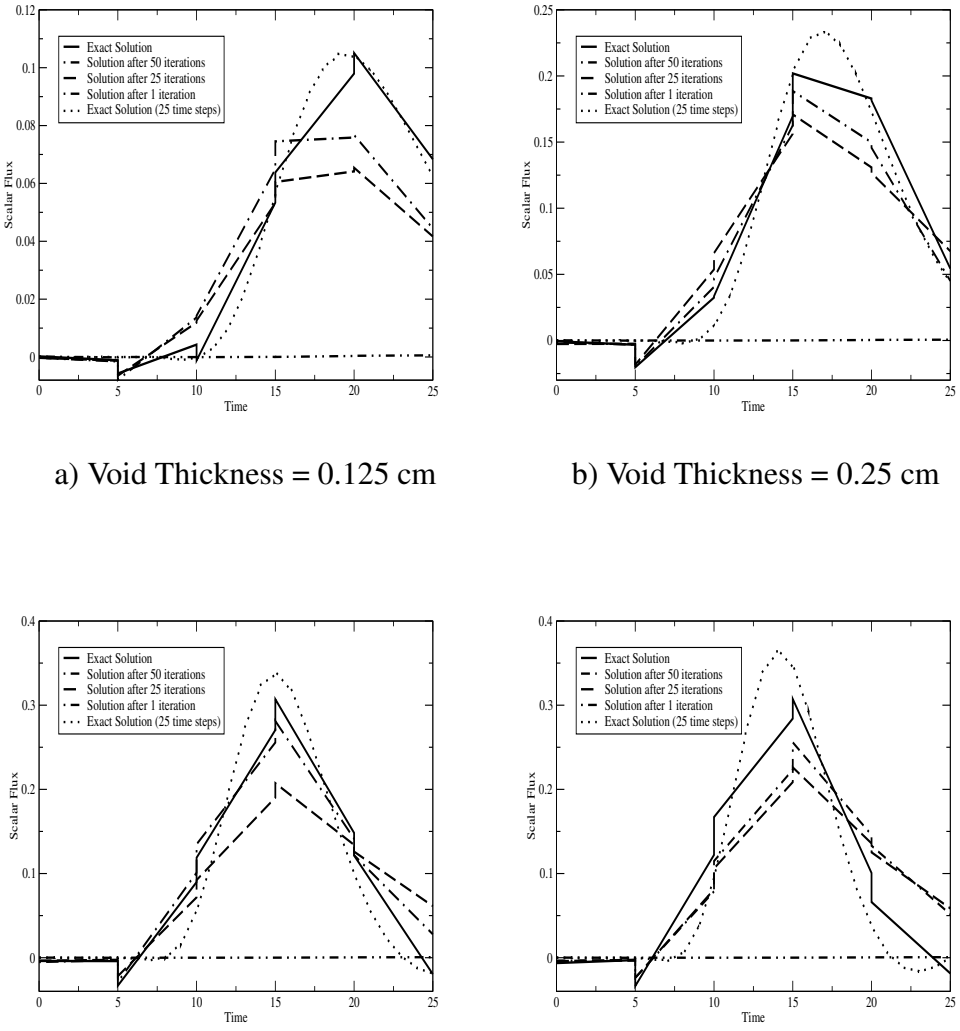


c) Detector 8

Figure 7.30: Scalar flux versus time for the first source at a) the sixth detector location, b) the seventh detector location, and c) the eighth detector location. The graphs compare the exact solution (solid line) to the forward model result after 1 iteration (dot-dashed line) and the forward model result after 50 iterations (dashed line). Calculations were performed for a void of thickness 0.25cm. The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each graph.

elling domain and the convergence characteristics of the inverse problem. The similarity in scalar flux versus time graphs at the second and eighth detector locations, the third and seventh detector locations and the fourth and sixth detector locations indicate that the model is symmetric and that the sources and detectors have been placed equidistantly around the circumference of the medium. In other words, the second and eighth detectors, third and seventh detectors and the fourth and sixth detectors are placed at equal distances away from the source. Thus the signal received at each of the detector pairs remains the same. Another interesting feature observed during the second and eighth detector response is the peak in the scalar flux during the initial part of the time sequence. This peak is observed when calculations are performed for 25 Discontinuous Galerkin time steps and represents the short amount of time the source pulse is activated and the integration of the spatial resolution. In other words the peak is a numerical representation of the delta function in time. The source has been placed on the first of five Discontinuous Galerkin time units.

The convergence characteristics show that as the number of iterations is increased the detector response at each of the source-detector pairs approaches that of the exact solution. Figure 7.31 illustrates scalar flux as a function of time for the first source and the third detector position. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (two dots and a dashed line), the forward model result after 25 iterations (dashed line), and the forward model result after 50 iterations (dot dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each individual graph. The top left hand side graph corresponds to results obtained using a void thickness of 0.125cm while the top right hand side graph describes results obtained using a void thickness of 0.25cm. The bottom left hand side graph corresponds to results obtained using a void thickness of 0.5cm while the bottom right hand side graph describes results obtained using a void thickness of 0.75cm. After 50 iterations the scalar flux graphs at each of the source-detector pairs converges close to that of the exact solution. The convergence criterion is chosen so that once convergence is reached, increasing the number of iterations does not change the inversion model significantly.



a) Void Thickness = 0.125 cm

b) Void Thickness = 0.25 cm

c) Void Thickness = 0.50 cm

d) Void Thickness = 0.75 cm

Figure 7.31: Scalar flux versus time for the first source and the third detector position. The graph compares the exact solution (solid line) to the forward model result after 1 iteration (two dots and a dashed line), the forward model result after 25 iterations (dashed line), and the forward model result after 50 iterations (dot dashed line). The exact solutions using 25 Discontinuous Galerkin time steps are also illustrated (dotted line) in each individual graph. The top left hand side graph corresponds to results obtained using a void thickness of 0.125cm while the top right hand side graph describes results obtained using a void thickness of 0.25cm. The bottom left hand side graph corresponds to results obtained using a void thickness of 0.5cm while the bottom right hand side graph describes results obtained using a void thickness of 0.75cm.

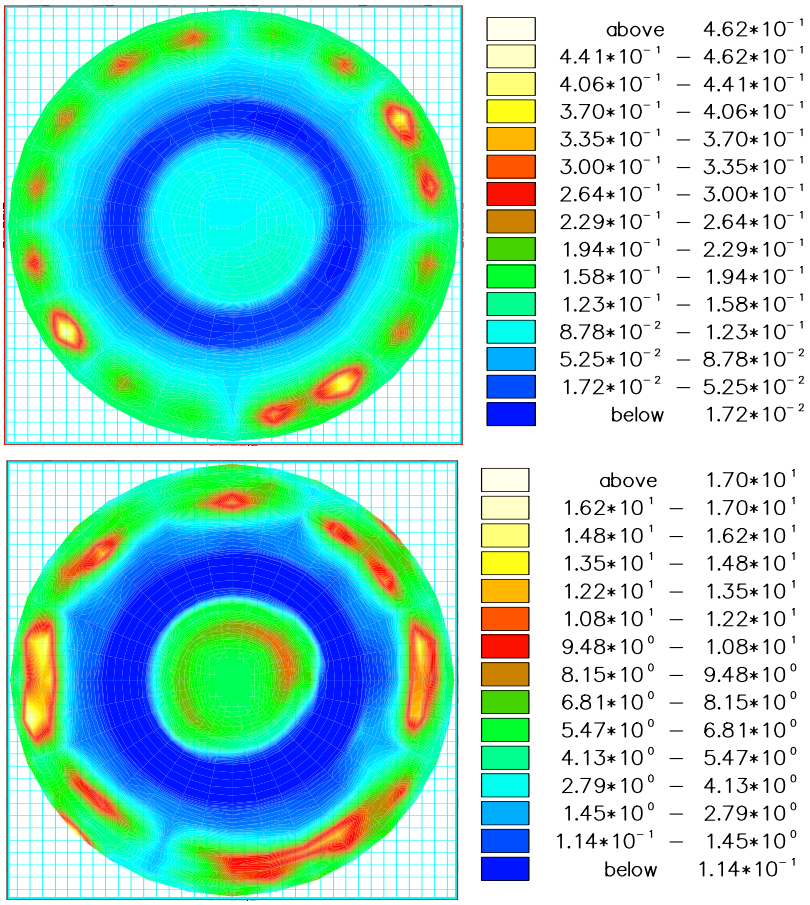


Figure 7.32: Absorption (top image) and scattering (bottom image) reconstruction results for a circular problem containing an embedded transparent ring. The ring is 0.5cm thick and is embedded 1cm deep inside the domain. The inversion was performed for 50 iterations using a P_1 angular expansion. The data contained 20 $\%$ random noise.

Incorporating Noise into the Time Dependent Imaging Domain

The ill-posed nature of inverse problems means that there are limitations on what images can accurately be reconstructed. The images also degrade when noise is added to the data. When developing a reconstuction algorithm it is usual to initially test that algorithm on simulated data before adding measurement error to the image reconstruction. A pseudo random number generator has been used to add independent Gaussian noise to each measurement. The independent Gaussian noise provides a good approximation to the true statistics of the data error [176].

Figure 7.32 illustrates the absorption (top image) and scattering (bottom image) reconstruction images for the circular problem described in figures 7.23 and 7.25. Figure 7.32 contains a 0.5cm thick transparent ring embedded 1cm deep inside the domain. The inversion was performed for 50 iterations using a a transient (Δt = 5) P_1 angular expansion. The data also contained 20 $\%$ random noise. The absorption and scattering coefficients of the background medium and the embedded void are the same as the absorption and scattering coefficients described in table 7.9.

The algorithm finds the correct location and approximate size of the embedded void for both absorption and scattering reconstructions despite the substantial noise added to the imaging domain. A slight degradation in image quality is observed but the additional information supplied by utilising time dependence minimises this effect.

7.4 Conclusion

A multi-dimensional inversion scheme has been developed to describe the transport of near infrared photons during medical optical tomography. The transport of optical radiation is described according to the one-speed Boltzmann transport equation. The discretisation has been designed specifically for inverse problems and is designed to work well for both optically thick (diffusive) and transparent media.

Four synthetic models have been used to demonstrate the simultaneous reconstruction of absorption and scattering material properties inside a host medium containing one or more transparent voids. While, the phantom problems are simple it is hoped that the inversion method will eventually be applied to real life medical imaging scenarios. Examples of transparent regions encountered in biological media include the cerebrospinal fluid that surrounds the human brain and synovial fluid contained between joints. The examples illustrate the enhanced imaging achievable when time dependent information is incorporated into the imaging domain and when full radiation transport theory is considered.

Chapter 8

Conclusion

Most pathological conditions are caused by a change in the biochemistry of tissue. Chemical changes can cause deficiencies in organ function and changes in the physical structure of tissue. The most reliable way of diagnosing and monitoring the treatment of disease is to look for changes in tissue samples that have been taken from a patient during a biopsy. Medical imaging methods provide a means of probing biological media, without the need for invasive surgery. Optical tomography provides a low cost alternative to conventional imaging tools such as X-ray imaging; X-ray computed tomography, ultrasonic imaging, magnetic resonance imaging and nuclear medicine. During Medical optical tomography near infrared radiation (of wavelength 650-900nm) is incident on the surface of a biological medium, in an attempt to reconstruct the spatial distribution of internal tissue optical properties. While optical tomography provides a low cost, non-invasive functional imaging method, a lack of appropriate instrumentation and the absence of sophisticated scattering models have limited its development as a conventional imaging tool.

For strongly scattering environments (such as those posed in near infrared optical tomography) the progressive influence of scattering with distances makes the scattering problem complex and non-linear. There is not a generally applicable direct method for imaging highly scattering phenomena and, instead, non-linear inversion methods are applied to evaluate optical properties. A multi-dimensional inversion technique has been developed here to describe the transport of photons during optical tomography. The nature of the

inverse problem is to determine the two and three dimensional distribution of material properties contained deep inside a host medium both for arbitrary model geometries and for arbitrary source and detector positions along the boundary of the domain. As such the flow of neutral particles through a medium of randomly scattering and absorbing particles is considered.

8.1 Radiation Transport Using the Boltzmann Transport Equation

The quality of optical property reconstruction is strongly dependent on the accuracy of the forward model. The propagation of light is fundamentally described according to electromagnetic theory but can be simplified to (radiation) transport theory by ignoring wave phenomena such as polarisation and interference and particle effects such as inelastic collisions. Transport theory provides a deterministic numerical method that describes the migration of neutral particles according to the Boltzmann transport equation. In this work both steady state and transient forms of the mono-energetic Boltzmann transport equation have been used to model the distribution of optical radiation inside a given medium. Prior to inversion a direct numerical discretisation scheme, subdivides the solution domain into a finite number of angle, ($\mathbf{\Omega}$ = (θ, ω), space $(\mathbf{r} = (x, y, z))$, and time, (t), components. Spatial discretisation is achieved using a Streamline Upwinded Petrov-Galerkin (SUPG) formulation, the angle of photon travel is described according to spherical harmonics and the time variable is discretised using a Discontinuous Galerkin (DG) representation in the time domain.

The SUPG method has been designed specifically for inversion regimes and yields a set of discrete equations that can be differentiated with respect to an object's material properties. This allows gradients to be formed as part of the inversion procedure. The SUPG method also incorporates an exponential stabilisation matrix, in order to provide the numerical dissipation necessary to model biological media. The exponential stabilisation matrix provides an adequate amount of weighting to model both optically thick (diffusive) and

transparent media, media with highly scattering and absorbing material properties and media with abruptly changing material properties.

Many time discretisation techniques result in poorly conditioned linear equations. The time dependent signal is often resolved using either a transform (e.g. a Laplace transform) or a truncated Fourier series expansion, which frequently require complex checkpointing algorithms or a large amount of computer memory. As such, many research groups prefer not to resolve the time dependent signal, and instead chose to map the governing equations into the frequency domain. The Discontinuous Galerkin discretisation of the time domain has the ability to use large time steps while still achieving a high order of accuracy in time. The discretisation is second order accurate at the end of each time level and third order accurate at the beginning of each time level. The Discontinuous Galerkin method therefore saves on computer memory and may therefore not require the use of checkpointing algorithms. Unlike Fourier analysis, Discontinuous Galerkin discretisations do not require the time domain to be periodic and surpress Gibbs oscillations that result from abruptly changing time signals. For each of the geometries considered throughout this thesis, the photon velocity has been set arbitrarily to 1 and just 5 Discontinuous Galerkin time units are used.

8.2 Inversion Described as a Functional Optimisation

The inversion has been posed as a functional optimisation problem [41, 69, 74, 75, 76]. Functional optimisation techniques employ a forward model that provides detector readings based on estimates of the distribution of optical properties inside the medium. The predicted detector readings are compared alongside observed data and are quantified using an error functional. The functional is minimised by updating the trial optical properties and performing new forward calculations with these updated optical properties. The material properties are iteratively updated using a gradient based optimisation method until the predicted data agrees with the detector readings. The final distribution of optical properties then provides a useful image.

Non-linear inversion schemes are frequently ill-posed. Using appropriate model covariance constraints, a suitable data weighting regime, and including any available *a priori* information can remedy concerns associated with non-uniqueness and limited data coverage. Functional optimisation techniques are highly desirable because they allow the straightforward inclusion of model covariance constraints by including additional terms in the error functional. Model covariance regularisation overcomes the problem of ill-posedness at the expense of limiting the allowed models to a class of model that is compatible with the provided model covariance information. Calculations performed throughout this thesis have incorporated regularisation terms associated with penalising both structure and deviation from the starting model.

During structural regularisation, large regularisation constraints are initially used that provide homogeneous model updates. The regularisation constraints are relaxed as iterations progress. This ensures large scale structures emerge during early iterations and that more complex structures emerge during later iterations. Regularisation associated with the deviation from the starting model is achieved by controlling the conditioning (i.e. the diagonal dominance) of the Hessian matrix to ensure it does not become singular and to ensure that the preconditioning matrix remains a good approximation of the Hessian matrix. Data weighting has also been used to assist the inversion. The weighting function typically has three contributions. In the first two contributions each source-detector pair is weighted according to the confidence in the datum and the preferential weighting of the datum. The final contribution arises from the way in which sources and detectors are distributed using a Gaussian approximation. Generally, data weighting has been chosen so as not to put excessive emphasis on short range inversion information obtained by placing sources and detectors close to one another.

Our research has primarily focused on the development of a gradient based, second order Levenberg-Marquardt optimisation scheme. The Levenberg-Marquardt method is a least squares optimisation technique in which matrix-vector multiplication between the first order gradient of the error functional and second order approximate Hessian matrix is sought. Although Levenberg-Marquardt methods involve computationally time consuming calculations involving both Jacobian and approximate Hessian matrices, the assembly and storage of the Jacobian and approximate Hessian matrix may be bypassed

alongside the need to implicitly invert the approximate Hessian. This is achieved using preconditioned linear conjugate gradient (PCG) solvers. Therefore a variant of the Levenberg-Marquardt optimisation strategy has been developed, in which only vector-vector multiplication of these matrices is required.

Another significant advantage of using second-order techniques such as the Levenberg-Marquardt method is its ability to treat regularisation terms implicitly. First order techniques, such as non-linear conjugate gradient methods, are guided primarily by the gradient of the error functional. Material properties are therefore mainly adjusted around the sources and detectors, since it is here that the gradient sensitivities are largest. During the first non-linear conjugate gradient iteration, the gradient contains no regularisation contribution since the initial model is homogenous. Therefore the model after the first non-linear conjugate gradient iteration is not influenced by regularisation and any damage caused by a step in a structurally unregularised direction remains throughout the inversion (this could provide and example of null space contamination [176])

. In contrast, the implicit nature of the Levenberg-Marquardt method enables structurally regularised inversion so that deeply embedded objects may be resolved even during the first Levenberg-Marquardt iteration. The regularisation parameters prevent material properties being adjusted primarily around the sources and detectors. Second order Levenberg-Marquardt calculations have been compared alongside first order non-linear conjugate gradient methods to illustrate the enhanced imaging achievable using implicit regularisation and second order inversion techniques.

8.3 Reconstruction Results

A number of synthetic problems are used to demonstrate the simultaneous reconstruction of absorption and scattering material properties inside a host medium. While, the phantom problems are simple it is hoped that the inversion method will eventually be applied to real life medical imaging scenarios. The examples illustrate the need to use full radiation transport theory, when considering transparent regions, during steady state

optical imaging. The examples also illustrate the enhanced imaging achievable when time dependent information is incorporated into the imaging domain.

Since the transport of radiation through biological tissue is often regarded as diffuse, many groups still use radiation transport models that rely on the diffusion approximation to the more generally applicable Boltzmann transport equation. The diffusion approximation represents a spherical harmonic expansion of the Boltzmann transport equation truncated to first order (i.e. a P_1 angular expansion). While diffusion theory remains a good approximation for light transport in tissue, diffusion theory fails in regions that are not regarded as diffuse, such as the clear, virtually transparent layer of cerebrospinal fluid (CSF) that circulates around the human brain or the synovial fluid contained between joints. These regions do not comply with the constraints of the diffusion approximation and must therefore be accounted for. Radiation transport models that rely on the Boltzmann transport equation can tolerate tissues that do not comply with the constraints of the diffusion approximation. Radiation transport models provide a suitable alternative to diffusion theory but are conceptually more difficult and require large computational time scales to perform an image reconstruction.

A full radiation transport model has been considered and compared alongside calculations performed using the diffusion approximation to the Boltzmann transport equations. Calculations have been performed using either P_1, P_3 or P_5 angular expansions. As the order of the angular approximation is increased, image resolution and contrast is generally improved. However, the computational time scales required to perform an image reconstruction are also increased. For strongly scattering environments a P_1 angular approximation (and hence diffusion theory) provided a suitable means of modelling radiation transport. For imaging domains containing transparent regions a P_1 angular approximation was not found to provide a suitable means for modelling radiation transport and instead both P_3 and P_5 angular expansions provided suitable alternatives. The P_1 angular expansion failed completely when trying to reconstructed embedded ring-like structures during steady state radiation transport.

The enhanced imaging achievable using time dependent information has also been considered. Time dependent radiation transport provides a significant improvement in image

quality and contrast during both P_1 and P_3 angular approximation reconstructions. The ability to use large time steps has enabled the analysis of the temporal signal without needing to significantly increase computer memory. The computational time scales required to perform an image reconstruction are also not prohibitive. An interesting observation is the significant improvement in image quality achieved when reconstructing transparent rings embedded into highly scattering environments. A P_1 angular approximation performed using time dependent radiation transport provided a significant improvement in image quality compared to both P_1 and P_3 steady state angular expansions. The transient P_1 inversion managed to reconstruct a ring-like structure for both absorption and scattering reconstructions, while the P_3 steady state approximation managed to resolve a ring-like structure for scattering coefficient reconstructions only. The P_1 steady state approximation was not able to resolve a ring-like structure during either the absorption coefficient or the scattering coefficient reconstructions.

8.4 Recommendations for Future Work

This thesis has presented a number of mathematical techniques that have contributed towards the development of optical tomographic image reconstruction. Future research must concentrate on improving the contrast and resolution of images, removing unwanted image artefacts and performing progressively more sophisticated and clinically relevant examples. A systematic investigation is therefore needed to fully understand the limitations of the reconstruction algorithm. Limiting factors that ought to be considered include geometry, the location of the sources and detectors and the appropriate use of *a priori* information. Specific suggestions for future research are further summarised:

Anisotropic Scattering and Anisotropic Materials

The reconstruction algorithm that has been described models a simplistic geometry in which both particle scattering and the medium's material properties are assumed to be isotropic. Biological tissue has an anisotropic, inhomogeneous material constitution [7,

64, 103], which can restrict photon transport in certain directions. In a layered tissue environment the subsurface is found to be smooth in the layering plane but rough perpendicular to the layering plane. It is therefore essential to model anisotropic media when considering particle scatter [177, 178, 179, 180, 181]. This is achieved by making the absorption and scattering coefficients angularly dependent ($\mu_a(\mathbf{r}, \mathbf{\Omega})$, and $\mu_s(\mathbf{r}, \mathbf{\Omega})$) and expanding that angular dependence into a finite set of spherical harmonic, P_N, functions.

Anisotropic scattering [182] in biological media is often described according to the Henyey-Greenstein function [46]. Anisotropic scattering is of particular interest when modelling largely voided configurations [179, 180]. If isotropic scattering is assumed, the small scattering and absorption coefficients associated with photon travel through transparent, void-like regions can be problematic especially when using low order S_N and P_N approximation techniques. It is therefore necessary to consider anisotropic scattering in regions where high levels of scattering are not assumed.

Anisotropic Regularisation

If the material properties are anisotropic, ambiguities associated with the angular dependent scattering anisotropy will be introduced into the domain. The regularisation matrix outlined in chapters 5, 6 and 7 should therefore be replaced with a more appropriate regularisation scheme [75, 183, 184]. It is hoped that an appropriate regularisation operator could be developed so that the optical tomographic inversion calculations provide unique solutions.

Geometry

Each of the examples presented in this study have been formulated as simple two dimensional geometric approximations of biological media. More complex and realistic geometries should be considered. It has been proposed that optical tomography could eventually be used in conjunction with another imaging technique to monitor the progression of disease. It would therefore be very useful to obtain some experimental data (for

example MRI or CT scans) so that the approximate geometry of a given tissue region may be taken into account. It is also necessary to model three dimensional data.

Experimental Data

The optical properties of biological tissues and the refractive index mismatches between boundary regions are largely unknown and therefore require further experimental investigation if realistic phantoms are to be considered.

Activatable Tracers

During near infrared imaging information is often lost because near infrared photons are scattered a large number of times before exiting the host medium. Fluorescent lifetime inversion [102, 185, 186] could remedy the problem of multiple scattering by introducing tracers into the imaging domain. Tracers are selected so that the photons emitted by the tracers, when incident with near infrared radiation, have wavelengths that are not as susceptible to scattering as near infrared wavelengths. This is equivalent to using contrasting agents in x-ray imaging, ultrasonic imaging and magnetic resonance imaging, or radiopharmaceuticals in nuclear medicine. The tracers are chosen according to their 'uptake' characteristics in certain tissue regions, organs or systems.

A multi-group [112] framework could be used to implement tracers or a method similar to the method developed by Hielscher et al [47, 187, 188] could be applied. Further optical fluorescence techniques are illustrated in [189, 190]. To model time-dependence, a delayed neutron framework [112] may be used to measure the time delay associated with the half-life of the florescent radiation emitted from transient particles.

Further Applications

While this thesis concentrates on medical imaging applications (and specifically medical optical tomography applications) a range of further applications could also be consid-

ered. Subject areas that could benefit from inversion technology include oceanography, atmosperic science, astrophysics, geophysics and neutron physics.

Bibliography

[1] Schematic of the 32-channel Time-Resolved Imaging System at UCL London. *http://www.medphys.ucl.ac.uk/research/borg/research/monstir/technical.htm*.

[2] J. C. Hebden, S. R. Arridge, and D. T. Delphy. Optical Imaging in Medicine: I. Experimental Techniques. *Physics in Medicine and Biology*, **42**:825–840, (1997).

[3] S. R. Arridge and J. C. Hebden. Optical Imaging in Medicine II: Modelling and Reconstruction. *Physics in Medicine and Biology*, **42**:841–853, (1997).

[4] S. R. Arridge. 'Optical Tomography in Medical Imaging', Topical Review. *Inverse Problems*, **15**:R41–R93, (1999).

[5] A. P. Gibson, J. C. Hebden, and S. R. Arridge. Recent Advances in Diffuse Optical Imaging. *Physics in Medicine and Biology*, **50**:R1–R45, (2005).

[6] F. E. W. Schmidt, M. E. Fry, J. C. Hebden, and D. T. Delphy. A 32-Channel Time-Resolved Instrument For Medical Optical Tomography. *Review of Scientific Instruments*, **71**:256–265, (2000).

[7] F. E. W. Schmidt. *Development of a Time-Resolved Optical Tomography System for Neonatal Brain Imaging*. PhD thesis, University College London, London, UK, (1999).

[8] P. Cook, N. Trasi, C. R. E. de Oliveira, and J. Haigh. *Modelling Radiation Transport Exchange Within and Between 3-D Cloud Structures*. 18^{th} International Conference of Transport Theory, Rio de Janeiro, RJ, Brazil, July 20-25, (2003).

[9] N. J. McCormick. *Analytic Inverse Transport Algorithms*. 18^{th} International Conference of Transport Theory, Rio de Janeiro, RJ, Brazil, July 20-25, (2003).

[10] H. F. Campos-Velho, F. M. Ramos, E. S. Chalhoub, S. Stephany, J. C. Carvalho, and N. J. Ferreira. *Inverse Photon Transport: Space Science Applications*. 18^{th} International Conference of Transport Theory, Rio de Janeiro, RJ, Brazil, July 20-25, (2003).

[11] A. Belleni-Morante, R. Monaco, R. Riganti, and F. Salvarani. A numerical Approach for Inverse Problems in Photon Transport Inside an Interstellar Cloud. *Applied Mathematics and Computation*, **154**:115–126, (2004).

[12] P. French. Biomedical Optics. *Physics World*, pages 41–46, (1999).

[13] C. M. C. Tempany and B. J. McNeil. Advances in Biomedical Imaging. *American Medical Association*, **285**:562–567, (2001).

[14] E. M. Sevick-Muraca, C. L. Hutchinson, and D. Y. Paithankar. Optical Tissue Biodiagnostics Using Fluoresence Lifetime. *Optics and Photonics News*, pages 25–28, (1996).

[15] Y. Hoshi and M. Tamura. Dynamic Multichannel Near-Infrared Optical Imaging of Human Brain Activity. *Journal of Applied Physiology*, **75**:1842–1846, (1993).

[16] E. Gratton, V. Toronov, U. Wolf, M. Wolf, and A. Webb. Measurement of Brain Activity by Near-Infrared Light. *Journal of Biomedical Optics*, **10(1)**:011008–1 –011008–13, (2005).

[17] H. Kawaguchi, T. Hayashi, T. Kato, and E. Okada. Theoretical Evaluation of Accuracy in Position and Size of Brain Activity Obtained by Near Infrared Topography. *Physics in Medicine and Biology*, **49**:2753–2765, (2004).

[18] E. Okada, M. Firbank, M. Schweiger, S. R. Arridge, M. Cope, and D.T. Delphy. Theoretical and Experimental Investigation of Near Infrared Light Propagation in a Model of the Adult Head. *Applied Optics*, **36**:21–31, (1997).

[19] M. Cope. *The Application of Near-Infrared Spectroscopy to Non-Invasive Monitoring of Cerebral Oxygenation in the Newborn Infant*. PhD thesis, University College, London, UK, (1991).

[20] J. C. Hebden, A. Gibson, T. Austin, R. Md Yusof, N. Everdell, D. T. Delphy, S. R. Arridge, J. H. Meek, and J. S. Wyatt. Imaging Changes in Blood Volume and Oxygenation in the Newborn Infant Brain Using Three-Dimensional Optical Tomography. *Physics in Medicine and Biology*, **49**:1117–1130, (2004).

[21] A. Y. Bluestone, G. Abdoulaev, C. H. Schmitz, R. L. Barbour, and A. H. Hielscher. Three Dimensional Optical Tomography of Hemodynamics in the Human Head. *Optics Express*, **9**:272–286, (2000).

[22] J. C. Hebden, A. Gibson, R. Md Yusof, N. Everdell, E. M. C. Hillman, D. T. Delphy, S. R. Arridge, T. Austin, J. H. Meek, and J. S. Wyatt. Three-Dimensional Optical Tomography of the Premature Infant Brain. *Physics in Medicine and Biology*, **47**:4144–4166, (2002).

[23] J. Beuthan, U. Netz, O. Minet, A. D. Klose, A. H. Hielscher, A. Scheel, J. Henninger, and G. Müller. Light Scattering Study of Rheumatoid Arthritis. *Quantum Electronics*, **32**:945–952, (2002).

[24] A. H. Hielscher, A. D. Klose, A. K. Scheel, B. Moa-Anderson, M. Backhaus, U. Netz, and J. Beuthan. Sagittal Laser Optical Tomography for Imaging of Rheumatoid Finger Joints. *Physics in Medicine and Biology*, **49**:1147–1163, (2004).

[25] A. J. Fallgatter, M. Roesler, L. Sitzmann, A. Heidrich, T. J. Mueller, and W. K. Strik. Loss of Functional Hemispheric Asymmetry in Alzheimer's Dementia Assesed with Near-Infrared Spectroscopy. *Cognitive Brain Research*, **6**:67–72, (1997).

[26] Y. Xu, Q. Zhang, and H. Jiang. Optical Image Reconstruction of Non-Scattering and Low Scattering Heterogeneities in Turbid Media Based on the Diffusion Approximation Model. *Journal of Optics A: Pure and Applied Optics*, **6**:29–35, (2004).

[27] P. Gould. New Look at Breast Cancer. *Physics World*, pages 7–8, (2002).

[28] V. G. Peters, D. R. Wyman, M. S. Patterson, and G. L. Frank. Optical Properties of Normal and Diseased Human Breast Tissue in the Visible and Near-Infrared. *Physics in Medicine and Biology*, **35**:1317–1334, (1990).

[29] T. Dierkes, D. Grosenick, K. T. Moesta, M. Möller, P. M. Schlag, H. Rinneberg, and S. R. Arridge. Reconstruction of Optical Properties of Phantom and Breast Lesion in vivo from Paraxial Scanning Data. *Physics in Medicine and Biology*, **50**:2519–2542, (2005).

[30] A. Li, E. L. Miller, M. E. Kilmer, T. J. Brukilacchio, T. Chaves, J. Stott, Q. Zhang, T. Wu, M. Chorlton, R. H. Moore, D. B. Kopans, and D. A. Boas. Tomographic Optical Breast Imaging Guided by Three-Dimensional Mammography. *Applied Optics*, **42**:5181–5190, (2003).

[31] S. Fantini, S. A. Walker, M. A. Franceschini, M. Kaschke, P. M. Schlag, and K. T. Moesta. Assessment of the Size, Position, and Optical Properties of Breast Tumours in vivo by Noninvasive Optical Methods. *Applied Optics*, **37**:1982–1989, (1998).

[32] J. J. Duderstadt and W. R. Martin. *Transport Theory*. John Wiley and Sons, (1979).

[33] J. J. Duderstadt and L. J. Hamilton. *Nuclear Reactor Analysis*. John Wiley and Sons, (1976).

[34] G. C. Pomraning. Radiation Hydrodynamics. Technical report, Notes from a short course given at Los Alamos, (1982).

[35] S. R. Arridge and M. Schweiger. Image Reconstruction in Optical Tomography. *Philosophical Transactions of the Royal Society of London B*, **352**:717–726, (1997).

[36] B. Chen, K. Stamnes, and J. J. Stamnes. Validity of the Diffusion Approximation in Bio-Optical Imaging. *Applied Optics*, **40**:6356–6366, (2001).

[37] S. R. Arridge and W. R. B. Lionheart. Nonuniqueness in Diffusion-Based Optical Tomography. *Optics Letters*, **23**:882–884, (1998).

[38] C. R Simpson, M. Kohl, M. Essenpreis, and M. Cope. Near-Infrared Optical Properties of Ex Vivo Human Skin and Subcutaneous Tissue Measured Using the Monte Carlo Inversion Technique. *Physics in Medicine and Biology*, **43**:2465–2478, (1998).

[39] S. T. Flock, M. S. Patterson, B. C. Wilson, and D. R. Wyman. Monte Carlo Modeling of Light Propagation in Highly Scattering Tissues - I: Model Predictions and Comparison with Diffusion Theory. *IEEE Transactions on Biomedical Engineering*, **36**:1162–1168, (1989).

[40] A. H. Hielscher, G. Abdoulaev, A. D. Klose, A. Bluestone, and J. Lasker. *Medical Optical Tomography with the Equation of Radiative Transfer*. 18^{th} International Conference of Transport Theory, Rio de Janeiro, RJ, Brazil, July 20-25, (2003).

[41] A. D. Klose and A. H. Hieschler. Optical Tomography Using the Time-Independent Equation of Radiative Transfer - Part 2: Inverse Model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **72**:715–732, (2002).

[42] A. D. Klose and A. H. Hielscher. Quasi-Newton Methods in Optical Tomographic Reconstruction. *Inverse Problems*, **19**:387–409, (2003).

[43] V. L. Barnard, C. C. Pain, A. G. Buchan, M. D. Eaton, and A. J. H. Goddard. A Finite Element Optical Imaging Method for Optically Thick and Optically Thin Media. *In Preparation*, (2006).

[44] V. L. Barnard, C. C. Pain, A. G. Buchan, M. D. Eaton, and A. J. H. Goddard. A Finite Element Optical Imaging Method Using Time Dependent Information. *In Preparation*, (2006).

[45] A. D. Klose, U. Netz, J. Beuthan, and A. H. Hieschler. Optical Tomography Using the Time-Independent Equation of Radiative Transfer - Part 1: Forward Model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **72**:691–713, (2002).

[46] E. D. Aydin. *A Higher-Order Transport Model for Photon Propagation and its Applications to Optical Tomography*. PhD thesis, Imperial College, London, UK, (1999).

[47] A. D. Klose, V. Ntziachristos, and A. H. Hielscher. The Inverse Source Problem Based on the Radiative Transfer Equation in Optical Molecular Imaging. *Journal of Computational Physics*, **202**:323–345, (2005).

[48] E. D. Aydin, C. R. E. de Oliveira, and A. J. H. Goddard. A Comparison Between Transport and Diffusion Calculations Using A Finite Element-Spherical Harmonics Radiation Transport Method. *Medical Physics*, **29**:2013–2023, (2002).

[49] E. Okada, M. Schweiger, S. R. Arridge, M. Firbank, and D. T. Delphy. Experimental Validation of Monte-Carlo and Finite Element Methods for the Estimation of the Optical Path Length in Inhomogeneous Tissue. *Applied Optics*, **35**:3362–3371, (1996).

[50] O. Dorn. A Transport-Backtransport Method for Optical Tomography. *Inverse Problems*, **14**:1107–1130, (1998).

[51] O. Dorn. Scattering and Absorption Transport Sensitivity Functions for Optical Tomography. *Optics Express*, **7**:492–506, (2000).

[52] H. Dehghani, S. R. Arridge, M. Schweiger, and D. T. Delphy. Optical Tomography in the Presence of Void Regions. *Journal of the Optical Society of America A*, **17**:1659–1670, (2000).

[53] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, and J. P. Kaipio. Hybrid Radiative-Transfer-Diffusion Model for Optical Tomography. *Applied Optics*, **44**:876–886, (2004).

[54] J. Riley, H. Dehghani, M. Schweiger, S. R. Arridge, J. Ripoll, and M. Nieto-Vesperinas. 3D Optical Tomography in the Presence of Void Regions. *Optics Express*, **7**:462–467, (2000).

[55] G. Mitic, J. Kolzer, J. Otto, E. Plies, G. Solkner, and W. Zinth. Time Gated Transillumination of Biologoical Tissue and Tissue Like Phantoms. *Applied Optics*, **33**:6699–6710, (1994).

[56] J. C. Hebden and S. R. Arridge. Imaging Through Scattering Media by the use of an Analytical Model of Perturbation Amplitudes in the Time Domain. *Applied Optics*, **35**:6788–6796, (1996).

[57] M. Schweiger and S. R. Arridge. Direct Calculation with a Finite Element Method of the Laplace Transform of the Distribution of Photon Time of Flight in Tissue. *Applied Optics*, **36**:9042–9049, (1997).

[58] C. C. Pain, M. D. Eaton, R. P. Smedley-Stevenson, A. J. H. Goddard, M. D. Piggott, and C. R. E. de Oliveira. Streamline Upwind Petrov-Galerkin Methods for the Steady-State Boltzmann Transport Equation. *Computer Methods and Applied Mechanics and Engineering, in press*, (2005).

[59] M. D. Eaton. *A High Resolution Riemann Method for Solving Radiation Transport Problems on Unstructured Meshes*. PhD thesis, Imperial College, London, UK, (2004).

[60] M. D. Eaton, C. C. Pain, R. P. Smedley-Stevenson, A. J. H. Goddard, M. D. Piggott, and C. R. E. de Oliveira. Control-Volume Finite Element Methods: With Applications to the Boltzmann Transport Equation. *Submitted to Computer Methods and Applied Mechanics and Engineering*, (2005).

[61] C. C. Pain, M. D. Eaton, J. Bowsher, R. P. Smedley-Stevenson, A. P. Umpleby, C. R. E de Oliveira, and A. J. H. Goddard. Riemann Solvers on 3-D Unstructured Finite Element Meshes for Time Dependent and Steady-State Radiation Transport. *Submitted to Computer Methods and Applied Mechanics and Engineering*, (2005).

[62] M. D. Eaton, C. C. Pain, C. R. E. de Oliveira, and A. J. H. Goddard. *A high-order Riemann method for the Boltzmann transport equation*. American Nuclear Society, Nuclear Mathematical and Computational Sciences: A Century Anew, Galintburg, Tennessee, April 6-11, (2003).

[63] A. E Profio. Light Transport in Tissue. *Applied Optics*, **28**:2216–2222, (1989).

[64] W. F. Cheong, S. A. Prahl, and A. J. Welch. A Review of the Optical Properties of Biological Tissues. *IEEE Journal of Quantum Electronics*, **26**:2166–2185, (1990).

[65] C. F. Bohren and D. R. Huffman. *Absorption and Scattering of Light by Small Particles*. John Wiley and Sons, (1983).

[66] P. N. den Outer and T. Nieuwenhuizen. Location of Objects in Multiple-Scattering Media. *Journal of the Optical Society of America A*, **10**:1209–1218, (1993).

[67] R. L. Barbour, H. L. Graber, Y. Xu, Y. Pei, and R. Aronson. Strategies for Imaging Diffusing Media. *Transport Theory and Statistical Physics*, **33**:361–371, (2004).

[68] J. C. Hebden, F. E. W. Schmidt, M. E. Fry, M. Schweiger, E. M. C. Hillman, D. T. Delphy, and S. R. Arridge. Simultaneous Reconstruction of Absorption and Scattering Images by Multichannel Measurement of Purely Temporal Data. *Optics Letters*, **24**:534–536, (1999).

[69] S. R. Arridge and M. Schweiger. Gradient Based Optimisation Scheme for Optical Tomography. *Optics Express*, **2**:213–226, (1998).

[70] R. Roy and E. M. Sevick-Muraca. A Numerical Study of Gradient-Based Nonlinear Optimisation Methods for Contrast Enhanced Optics Tomography. *Optics Express*, **9**:49–65, (2001).

[71] N. Gershenfeld. *The Nature of Mathematical Modeling*. Cambridge University Press, (1999).

[72] J. Nocedal and S. J. Wright. *Numerical Optimisation*. Springer Series in Optimisation Research, Springer-Verlag, New York, Berlin, Heidelberg, (1999).

[73] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipies in Fortran 77: Volume. 1*. Cambridge University Press, Second Edition, (1992).

[74] C. C. Pain, J. V. Herwanger, and M. H. Worthington. Effective Multi-Dimensional Resistivity Inversion Using Finite Element Techniques. *Geophysical Journal International*, **151**:710–728, (2002).

[75] C. C. Pain, J. V. Herwanger, J. Saunders, M. H. Worthington, and C. R. E. de Oliveira. Anisotropic Resistivity Inversion. *Inverse Problems*, **19**:1081–1111, (2003).

[76] J. H. Herwanger. *Seismic and Electric Crosshole Tomography for Fracture Detection and Charaterisation*. PhD thesis, Imperial College, London, UK, (2001).

[77] A. Tarantola. *Inverse Problem Theory*. Elsevier Science, (1987).

[78] M.Bertero and P. Boccacci. *Introduction to Inverse Problems In Imaging*. Institute of Physics, (1998).

[79] I. M. Navon. *Adjoint Equations and Inverse CFD Problems*. in preparation, (2005).

[80] A. Webb. *An Introduction to Biomedical Imaging*. John Wiley and Sons, (2003).

[81] G. Michael. X-ray Computed Tomography. *IOP Publishing*, pages 442–451, (2001).

[82] J. Rydberg, K. A. Buckwalter, K. S. Caldemeyer, M. D. Phillips, D. J. Conces, A. M. Aisen, S. A. Persohn, and K.K. Kopecky. Multisection CT: Scanning Techniques and Clinical Applications. *Radiographics*, **20**:1787–1806, (2000).

[83] T. Fuchs, M. Kachelreiß, and W. A. Kalender. Technical Advances in Multi-Slice Spiral CT. *European Journal of Radiology*, **36**:69–73, (2000).

[84] M. Claudon, F. Tranquart, D. H. Evans, F. Lefevre, and J.M. Correas. Advances in Ultrasound. *European Journal of Radiology*, **12**:7–18, (2002).

[85] T. R. Nelson and D. H. Pretorius. Three-Dimensional Ultrasound Imaging. *Ultrasound in Medicine and Biology*, **24**:1243–1270, (1998).

[86] A. Hein and W. D. O'Brien. Current Time-Domain Methods for Assessing Tissue Motion by Analysis from Reflected Ultrasound Echoes-A Review. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, **40**:84–102, (1993).

[87] W. Becker and J. Meller. The Role of Nuclear Medicine in Infection and Inflammation. *The Lancet, Infectious Diseases*, **1**:326–333, (2001).

[88] M. W. Groch and W. D. Erwin. SPECT in the Year 2000: Basic Principles. *Journal of Nuclear Medicine Technology*, **28**:233–244, (2000).

[89] T. G. Turkington. Introduction to PET Instrumentation. *Journal of Nuclear Medicine Technology*, **29**:4–11, (2001).

[90] D. Halliday, R. Resnick, and J. Walker. *Fundamentals of Physics Extended*. John Wiley and Sons, (1997).

[91] M. L. Boas. *Mathematical Methods in the Physical Sciences*. John Wiley and Sons, (1983).

[92] K. A. Stroud. *Further Engineering Mathematics*. Macmillan Press LTD, Third Edition, (1996).

[93] E. M. C. Hillman. *Experimental and Theoretical Investigations of Near Infrared Tomographic Imaging Methods and Clinical Applications*. PhD thesis, University College, London, UK, (2002).

[94] V. S. Hollis. *Non-Invasive Montioring of Brain Tissue Temperature by Near Infrared Spectroscopy*. PhD thesis, University College, London, UK, (2002).

[95] A. Gibson, R. Md Yusof, H. Dehghani, J. Riley, N. Everdell, R. Richards, J. C. Hebden, M. Schweiger, S. R. Arridge, and D. T. Delphy. Optical Tomography of a Realistic Neonatal Head Phantom. *Applied Optics*, **42**:3109–3116, (2003).

[96] P. G. B. Johnston. *The Newborn Child*. 8^{th} Edition, New York, Churchill, Livingston, (1998).

[97] C. J. H. Kelnar, D. Harvey, and C. Simpson. *The Sick Newborn Baby*. 3^{rd} Edition, Philadelphia: Bailliere Tindal, (1995).

[98] G. B. Marenstein and S. L. Gardner. *Handbook of Neonatal Intensive Care*. 4^{th} Edition, St Louis London: Mosby, (1998).

[99] J. C. Hebden. Advances in Optical Imaging of the Newborn Infant Brain. *Psychophysiology*, **40**:501–510, (2003).

[100] R. Marchesini, C. Clemente, E. Brambilla, and M. Brambilla. Optical Properties of In Vitro Epidermis and their Possible Relationship with Optical Properties of In

Vivo Skin. *Journal of Photochemisty and Photobiology B: Biology*, **16**:127–140, (1992).

[101] B. Devaraj, M. Takeda, M. Kobayashi, M. Usa, K. P. Chan, Y. Watanabe, T. Yuasa, T. Akatsuka, M. Yamada, and H. Inaba. In Vivo Laser Computed Tomographic Imaging of Human Fingers by Coherent Detection Imaging Method Using Different Wavelengths in Near Infrared Region. *Applied Physics Letters*, **69**:3671–3673, (1997).

[102] B. P. Ölveczky. Fluoresence lifetime imaging (flim): Tomographic reconstruction of 3-d fluoresence lifetime distributions. Master's thesis, Imperial College, London, UK, (1996).

[103] S. Nickell, M. Hermann, M. Essenpreis, T. J. Farrell, U. Kramer, and M. S. Patterson. Anisotropy of Light Propagation in Human Skin. *Physics in Medicine and Biology*, **45**:2873–2886, (2000).

[104] H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson. Simultaneous Reconstruction of Optical Absorption and Scattering Maps in Turbid Media from Near Infrared Frequency-Domain Data. *Optics Letters*, **20**:2128–2130, (1995).

[105] M. Schweiger and S. R. Arridge. Optical Tomographic Reconstruction in a Complex Human Head Model Using A Priori Region Boundary Information. *Physics in Medicine and Biology*, **44**:2703–2721, (1999).

[106] C. Cecchi-Pestellini, L. Barletti, S. Aiello, and A. Bellini-Morante. Mathematical Methods for Photon Transport in Random Media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **65**:835–851, (2000).

[107] D. Contini. Human Capital and Mobility Project - Activity and Networks; Modelling Photon Migration in Turbid Media. Technical report, University of Florence and Imperial College, London, (1999).

[108] R. Nossal, J. Kiefer, G. H. Weiss, R. Bonner, H. Taitelbaum, and S. Havlin. Photon Migration in Layered Media . *Applied Optics*, **27**:3382–3391, (1988).

[109] M. F. Modest. *Radiative Heat Transfer*. Elsevier Science, (1993).

[110] M. M. R. Williams. *Mathematical Methods in Particle Transport Theory*. Butterworks, London, (1971).

[111] N. Metropolis and S. Ulam. The Monte Carlo Method. *Journal of the American Statistics Association*, **44**:335, (1949).

[112] E. E. Lewis and W. F. Miller Jr. *Computational Methods of Neutron Transport*. American Nuclear Science, (1993).

[113] S. A. Prahl, M. Keijzer, S. L. Jacques, and A. J. Welch. A Monte Carlo Model of Light Propagation in Tissue. *Applied Optics*, **IS 5**:102–111, (1989).

[114] D. A. Boas, J. P. Culver, J. J. Stott, and A. K. Dunn. Three Dimensional Monte Carlo Code for Photon Migration Through Complex Heterogeneous Media Including the Adult Human Head. *Optics Express*, **10**:159–170, (2002).

[115] A. H. Hieschler, R. E. Alcouffe, and R. L. Barbour. Comparison of Finite Difference Transport and Diffusion Calculations for Photon Migration in Homogeneous and Heterogeneous Tissues. *Physics in Medicine and Biology*, **43**:1285–1302, (1998).

[116] R. J. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, Cambridge, UK, (2002).

[117] D. Henwood and J. Bonet. *Finite Elements: A Gentle Introduction*. Macmillan Press Limited, Basingstoke, Hampshire and London, (1996).

[118] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method, Volume 1: The Basis, Fifth Edition*. Butterworth Heinmann, Oxford, UK, (2000).

[119] K. F. Hansen and C. M. Kang. Finite Element Methods in Reactor Physics Analysis. *Advances in Nuclear Science and Technology*, **8**:175–253, (1975).

[120] S. Sherwin and I. Matthews. Finite Element Methods. Technical report, Imperial College, London, (2003).

[121] A. G. Buchan. Anglular Discretisation of the First Order Boltzmann Transport Equation. Technical report, Student Report, Imperial College, London, UK, (2003).

[122] A. G. Buchan. Anglular Discretisation of the First Order Boltzmann Transport Equation, Part 4: P_N, SP_N and Diffusion Methods. Technical report, Imperial College, London, UK, (2003).

[123] S. L. Jacques. Time Resolved Propagation of Ultrashort Laser Pulses Within Turbid Tissues. *Applied Optics*, **28**:2223–2229, (1989).

[124] A. Sood and E. Selcow. MCNP Intermediate Training Course. Technical report, Notes from a Short Course Taken at Imperial College, London, (2003).

[125] C. R. E. de Oliveira. *Finite Element Techniques for Multigroup Neutron Transport Calculations with Anisotropic Scattering*. PhD thesis, University of London, Queen Mary's College, London, UK, (1983).

[126] R. K. Nesbet. *Variational Principles and Methods in Theoretical Physics and Chemistry*. Cambridge University Press, Cambridge, (2003).

[127] J.E. Morel and J.M. McGhee. A Self-Adjoint Angular Flux Equation. *Nuclear Science and Engineering*, **132**:312, (1999).

[128] S. R. Arridge and M. Schweiger. Photon Measurement Density Functions Part 2: Finite Element Method Calculations. *Applied Optics*, **34**:8026–8037, (1995).

[129] C. C. Pain, M. D. Eaton, R. P. Smedley-Stevenson, A. J. H. Goddard, M. D. Piggott, and C. R. E. de Oliveira. Space-Time Streamline Upwind Petrov-Galerkin Methods for the Boltzmann Transport Equation. *Computer Methods and Applied Mechanics and Engineering, in press*, (2005).

[130] B .W. Pogue and M. S. Patterson. Forward and Inverse Calculations for Near Infra-Red Imaging Using a Multigrid Finite Difference Method. *OSA Proceedings on Advances in Optical Imaging and Photon Migration*, **21**:176–180, (1994).

[131] A. Ishmariu. *Wave Propagation and Scattering in Random Media: Vol. 1*. Academic Press, (1978).

[132] H. J. G. M. Sterenborg, M. J. C. Van-Gemert, W. Kamphorst, and J. G. Wolbers. The Spectral Dependence of Optical Properties of the Human Brain. *Lasers in Medical Science*, **4**:221–227, (1989).

[133] P. van der Zee. *Measurement and Modelling of the Optical Properties of Human Tissue in the Near Infrared*. PhD thesis, University College, London, UK, (1993).

[134] F. Bevilacqua, D. Piguet, P. Marquet, J. D. Gross, B.J. Tromberg, and C. Depeursinge. Local Determination of Tissue Optical Properties: Applications to Human Brain. *Applied Optics*, **38**:4939–4950, (1999).

[135] G. M. Hale and M. R. Querry. Optical Constants of Water in the 200nm-200 μm Wavelength Region. *Applied Optics*, **12**:555–563, (1973).

[136] M. Firbank, M. Hiraoka, M. Essenpreis, and D. T. Delphy. Measurement of the Optical Properties of the Skull in the Wavelength Range 650-950nm. *Physics in Medicine and Biology*, **38**:503–510, (1993).

[137] J. R. Mourant, J. P. Freyer, A. H. Hielscher, A. A. Eick, D. Shen, and T. M. Johnson. Mechanisms of Light Scattering from Biological Cells Relevant to Noninvasive Optical-Tissue Diagnostics. *Applied Optics*, **37**:3586–3593, (1998).

[138] B. C. Wilson and S. L. Jacques. Optical Reflectance and Transmittance of Tissues: Principles and Applications. *IEEE Journal of Quantum Electronics*, **26**:2166–2185, (1990).

[139] M. R. Arnfield, R. P. Mathew, J. Tulip, and M. S. McPhee. Analysis of Tissue Optical Coefficients Using an Approximate Equation Valid for Comparable Absorption and Scattering. *Physics in Medicine and Biology*, **37**:1219–1230, (1992).

[140] J. F. Beek, P. Blokland, P. Posthumus, M. Aalders, J. W. Pickering, H. J. C. M. Sterenborg, and M. C. J. van Gemet. In Vitro Double-Integrating-Sphere Optical Properties of Tissues Between 630 and 1064nm. *Physics in Medicine and Biology*, **42**:2255–2261, (1997).

[141] S. Fantini, D. Hueber, M. A. Franceschini, E. Gratton, W. Rosenfeld, P. G. Stubblefield, D. Maulik, and M. R. Stankovic. Non-Invasive Optical Monitoring of

the Newborn Piglet Brain Using Continuous-Wave and Frequency-Domain Spectroscopy. *Physics in Medicine and Biology*, **44**:1543–1563, (1999).

[142] R. M. P. Doornbos, R. Lang, M. C. Aalders, F. W. Cross, and H. J. C. M. Sterenborg. The Determination of In Vivo Human Tissue Optical Properties and Absolute Chromophore Concentrations Using Spatially Resolved Steady-State Diffuse Reflectance Spectroscopy. *Physics in Medicine and Biology*, **44**:967–981, (1999).

[143] A. Torricelli, A. Pifferi, P. Taroni, E. Giambattistelli, and R. Cubeddu. In Vivo Optical Characterization of Human Tissues from 610-1010nm by Time-Resolved Reflectance Spectroscopy. *Physics in Medicine and Biology*, **46**:2227–2237, (2001).

[144] S. Fantini, M. A. Franceschini, E. Gratton, D. Heuber, W. Rosenfeld, D. Maulik, P. G. Stubblefield, and M. R. Stankovic. Non-Invasive Optical Mapping of the Piglet Brain in Real Time. *Optics Express*, **4**:308–314, (1999).

[145] E. Okada and D. T. Delphy. Near-Infrared Light Propagation in an Adult Head Model I. Modeling of Low-Level Scattering in the Cerebrospinal Fluid Layer. *Applied Optics*, **42**:2906–2914, (2003).

[146] D. Parsons-Karavassilis. Fabrication and characterisation of optical phantoms for biomedical imaging. Master's thesis, Imperial College, London, UK, (1997).

[147] F. X. le Dimet, I. M. Navon, and D. N. Dacseu. Second Order Information in Data Assimilation. *Monthly Weather Review*, **130**:629–648, (2001).

[148] Y. Zhu and I. M. Navon. Documentation of the Tangent Linear and Adjoint Models of the FSU Global Spectral Model T42L12. Technical report, Supercomputer Computations Research Institute - FSU, (1998).

[149] R. L. Burden, J. D. Faires, and A. C. Reynolds. *Numerical Analysis*. 2^{nd} Edition: Prindle, Weber and Schmidt: Boston, Massachusetts, (1981).

[150] F. Fang. Inversion of the Linear 1D Water Wave Equation. Technical report, Imperial College, London, UK, (2002).

[151] Y. Zhu and I. M. Navon. Impact of Parameter Estimation on the Performance of the FSU Global Spectral Model Using Its Full Physics Adjoint. *Monthly Weather Review*, **127**:1497–1517, (1999).

[152] A. M. Moore. Data Assimilation in a Quasi-Geostrophic Open-Ocean Model of the Gulf Stream Region Using the Adjoint Method. *Journal of Physical Oceanography*, **21**:398–427, (1991).

[153] B. F. Sanders and N. D. Katopodes. Adjoint Sensitivity Analysis for Shallow-Water Wave Control. *Journal of Engineering Mechanics*, pages 909–919, (2000).

[154] R. Giering. Tangent Linear and Adjoint Biogeochemical Models. *Inverse Methods in Global Biogeochemical Cycles, Geophysical Monograph*, **114**:33–46, (2000).

[155] I. M. Navon. Variational Data Assimilation, Optimal Parameter Estimation and Sensitivity Analysis for Environmental Problems. Technical report, Lecture Notes, ICES, (1995).

[156] Y. Zhu and I. M. Navon. FSU-GSM Forecast Error Sensitivity in Initial Conditions: Applications to Indian Summer Monsoon. *Meterology and Atmospheric Physics*, **68**:35–41, (1998).

[157] S. K. Nadarajah, A. Jameson, and J. J. Alonso. An Adjoint Model for the Calculation of Non Collocated Sensitivity in Supersonic Flow. Technical report, 40^{th} AIAA Aerospace Meeting and Exhibit, Reno, Nevada, (2002).

[158] M. Schweiger, S. R. Arridge, and I. Nissila. Gauss-Newton Method for Image Reconstruction in Diffuse Optical Tomography. *Physics in Medicine and Biology*, **50**:2365–2386, (2005).

[159] A. H. Hielscher, A. D. Klose, and K. M. Hanson. Gradient-Based Iterative Image Reconstruction Scheme for Time-Resolved Optical Tomography. *IEEE Transactions on Medical Imaging*, **18**:262–271, (1999).

[160] J. Zhou and J. Bai. Spatial Location Weighted Optimization Scheme for DC Optical Tomography. *Optics Express*, **11**:141–150, (2003).

[161] I. Kwee. *Towards a Bayesian Framework for Optical Tomography*. PhD thesis, University College, London, UK, (1999).

[162] R. Giering. Recipies for Adjoint Code Construction. *ACM Transactions on Mathematical Software*, **24**:437–474, (1998).

[163] A. K. Alekseev and I. M. Navon. The Analysis of an Ill-Posed Problem Using Multi-Scale Resolution and Second Order Adjoint Techniques. *Comput. Methods. Appl. Mech. Eng*, **190**:1937–1953, (2001).

[164] A. G. Buchan, C. C. Pain, M. D. Eaton, R. P. Smedley-Stevenson, A. J. H. Goddard, , and C. R. E. de Oliveira. Hierarchical Preconditioning for the Boltzmann transport Equation. *In Preparation*, (2005).

[165] T .M .R. Ellis, I. R. Philips, and T. M. Lahey. *Fortran 90 Programming*. Addison-Wesley, (1998).

[166] B. D. Hahn. *Fortran 90 for Scientists and Engineers*. Arnold, (1998).

[167] R. G. Pratt, C. Shin, and G. J. Hicks. Gauss-Newton and Full-Newton Methods in Frequency-Space Seismic Waveform Inversion. *Geophysical Journal International*, **133**:341–362, (1998).

[168] Y. Jianchao and C. T. Chern. *Comparison of Newton-Gauss with Levenberg-Marquardt Algorithm for Space Resection*. Paper Presented at the 22^{nd} Asian Conference on Remote Sensing, 5-9 November, (2001).

[169] A. Adler and W. R. B. Lionheart. Uses and Abuses of EIDORS: an Extensible Software Base for EIT. *Physiological Measurement*, **27**:S25–S42, (2006).

[170] S. J. Matcher. Nonuniqueness in Optical Tomography: Relevance of the P_1 Approximation. *Optics Letters*, **24**:1729–1731, (1999afhiklns).

[171] A. Griewank and A. Walther. Algorithm 799: Revolve: An implementation of Checkpointing for the Reverse Adjoint Mode of Computational Differentiation. *ACM Transactions on Mathematical Software*, **26**:19–45, (2000).

[172] J. B. Fishkin, O. Coquoz, E. R. Anderson, M. Brenner, and B. J. Tromberg. Frequency-Domain Photon Migration Measurements of Normal and Diseased Malignant Tissue Optical Properties in a Human Subject. *Applied Optics*, **36**:10–20.

[173] R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini. Imaging with Diffusing Light: An Experimental Study of the Effect of Background Optical Properties. *Applied Optics*, **37**:3564–3573, (1998).

[174] M. Keijzer, W. M. Star, and P. R. M. Storchi. Optical Diffusion in Layered Media. *Applied Optics*, **27**:1820–1824, (1988).

[175] J. Ripoll, M Nieto-Vesperinas, S. R. Arridge, and H. Dehghani. Boundary Conditions for Light Propagation in Diffusive Media with Nonscattering Regions. *Journal of the Optical Society of America A*, **17**:1671–1681, (2000).

[176] W. R. B. Lionheart. EIT Reconstruction Algorithms: Pitfalls, Challenges and Recent Developments. *Physiological Measurement*, **25**:125–142, (2004).

[177] J. Heino, S. R. Arridge, J. Sikora, and E. Somersalo. Anisotropic Effects in Highly Scattering Media. *Physical Review E*, **68**:031908–1 – 031908–8, (2003).

[178] V. I. Haltrin. Exact Solution of the Characteristic Equation of Transfer in the Anisotropically Scattering and Absorbing Medium. *Applied Optics*, **27**:599–602, (1988).

[179] M. M. R. Williams. Transport Theory in Anisotropic Media. *Mathematical Proceedings of the Cambridge Philosophical Society*, **84**:549–567, (1978).

[180] G. Lapenta, P. Ravetto, M. M. Rostagno, R. Jacqmin, F. Malvagi, and G. Rimpault. Neutron Transport Problems in Anisotropic Media. *Annals of Nuclear Energy*, **28**:1271–1286, (2001).

[181] C. Yildiz. Influence of Anisotropic Scattering on the Size of Time-Dependent Systems in Monoenergetic Neutron Transport. *Journal of Physics D: Applied Physics*, **32**:317–325, (1999).

[182] J. R. Mourant, J. Boyer, A. H. Hielscher, and I. J. Bigio. Influence of Scattering Phase Function on Light Transport Measurements in Turbid Media Performed with Small Source-Detector Separations. *Optics Letters*, **21**:546–548, (1996).

[183] J. V. Herwanger, C. C. Pain, and C. R. E. de Oliveira. *Electric and Seismic Inversion in Anisotropic and Inhomogeneous Media*. 4^{th} International Conference on Inverse Problems in Engineering, Rio de Janeiro, Brazil, (2002).

[184] A. H. Hielscher and S. Bartel. Use of Penalty Terms in Gradient-Based Iterative Reconstruction Schemes for Optical Tomography. *Journal of Biomedical Optics*, **6(2)**:183–192, (2001).

[185] A. J. Hughes. Fluorescence Lifetime Imaging with Application to in-vivo Diagnostics. Technical report, Imperial College. London, UK.

[186] A. Godavarty, A. B. Thompson, R. Roy, M. Gurfinkel, M. J. Eppstein, C. Zhang, and E. M. Sevick-Muraca. Diagnostic Imaging of Breast Cancer using Fluoresence-Enhanced Optical Tomography: Phatom Studies. *Journal of Biomedical Optics*, **9**:488–496, (2004).

[187] A. D. Klose and A. H. Hielscher. Fluorescence Tomography with Simulated Data Based on the Equation of Radiative Transfer. *Optics Letters*, **28**:1019–1021, (2003).

[188] A. D. Klose and A. H. Hielscher. *Fluorescence Tomography in Highly Scattering Media with the Equation of Radiative Transfer*. 18^{th} International Conference of Transport Theory, Rio de Janeiro, RJ, Brazil, July 20-25, (2003).

[189] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Milane. Fluorescence Optical Diffusion Tomography. *Applied Optics*, **42**:3081–3094, (2003).

[190] R. Roy, A. Godavarty, and E. M. Sevick-Muraca. Fluorescence-Enhanced Optical Tomography Using Referenced Measurements of Heterogeneous Media. *IEEE Transactions on Medical Imaging*, **22**:824–836, (2003).